# Structural characterization of components of the flagellar export apparatus from the *H. pylori*.

A Thesis Submitted to the College of

Graduate Studies and Research

In Partial Fulfillment of the Requirements

For the Degree of Masters of Science

In the Department of Biochemistry

University of Saskatchewan

Saskatoon

By

Dakshata Ambilwade

2016

**Permission to Use**

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or part, for scholarly purposes may be granted by the professors who supervised my thesis work, or in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and the University of Saskatchewan in any scholarly use which may be made of any materials in my thesis. Request for permission to copy or make other use of material in this thesis in whole or in part should be addressed to:

Head of the Department of Biochemistry
University of Saskatchewan
Saskatoon, Saskatchewan, S7N 5E5

# Abstract

*H. pylori* is a human gastric pathogen responsible for serious health conditions such as gastritis and peptic ulcers. It requires polar sheathed flagella for initial colonization and manifestation of infection in the human stomach. The export of flagellum components requires a specialized protein complex called the export apparatus. FliI ATPase, FliH, and FliJ are major soluble components of the flagellar export apparatus. FliI is thought to provide energy for export of flagellum components by converting the energy of ATP hydrolysis into energy for export of proteins. *In vitro,* FliH is a negative regulator of FliI ATPase activity. FliJ is a general export co-chaperone. In the case of *H. pylori*, there is no strong evidence for the existence of FliJ. There are other important proteins involved in flagellum assembly such as flagellin, FliS and FlhA. The project involves protein characterization and crystallization studies of proteins FliH, FliI, FliS, flagellin and the identification of FliJ in *H. pylori.*

The *in-vitro* characterization studies of two constructs of *H. pylori* FliH (FliH (57-258) and FliH (73-258)) suggests that FliH likely forms a homodimer in solution and it forms an elongated structure. The crystallization studies of both these proteins yielded beautiful crystals but getting structural information was challenging due to diffraction quality of the crystals. The site-directed mutagenesis of FliI was successfully performed to produce FliIE193Q mutant. The protein characterization studies showed that FliIE193Q mutant precipitates at every stage of purification. The gel filtration suggests that it forms a monomer in solution. The crystallization studies of FliIE193Q mutant did not produce any crystals possibly due to solubility issues. The interaction studies of FliH and the FliIE193Q mutant suggests that it most likely forms $FliH_2$:FliI complex.

Bioinformatics studies suggests HP0256 as a potential homolog of FliJ (Douillard *et al*., 2010). The protein characterization studies showed that HP0256 forms a monomer in solution. Unfortunately, no crystals were observed for HP0256. The structural studies of FliS showed that it forms an antiparallel four-helix bundle. The binding studies (GST-pulldown) of FliS and flagellin showed that the flagellin forms a complex with the FliS.

Overall, we have characterized some of the key proteins from the *H. pylori* flagellar export system.

# Acknowledgement

I would like to use this opportunity to express my gratitude to Dr. Stanley Moore for his aspiring guidance and supervision throughout my Masters project. I would also like to thank him for providing me lab space and other necessary resources for this research work.

I would like to thank my committee members, Dr. Jeremy Lee, Michel Fodje. Their timely supervision and friendly advice really helped me in this project.

I also thank Yunhua Jia (research associate) from Dr. Moore's lab for hands-on support throughout my lab work. I also wish to thank members of Moore lab for their support throughout this work: Sifa Arrafi and Poonam Dhindwal.

I would like to dedicate this thesis to Krishna, Mom, Dad, my family, all of whom always loved and supported me. Finally, thank you GOD.

# Table of Contents

# List of chemicals

**List of reagents, kits, instruments and Suppliers**

| Reagent list | Suppliers |
|---|---|
| Al's oil | Hampton Research, Aliso Viejo, CA |
| Ampicillin | Bioshop, Burlington, ON |
| AMP-PNP | Sigma-Aldrich, MO |
| BICINE | Sigma-Aldrich, MO |
| CAPS (N-cyclohexyl-3-aminopropanesulfonic acid) | Sigma-Aldrich, MO |
| CHES (*N*-Cyclohexyl-2-aminoethanesulfonic acid) | Bioshop, Berlignton, ON |
| Dialysis tubing (3000-8000 MWCO) | Biodesign Dialysis tubing™, NY |
| DTT | Bioshop, Berlignton, ON |
| Glucose Nutrient Mix | Athena Enzyme Systems™ |
| Glutathione sepharose™ 4-B | GE healthcare, Uppsala, Sweden |
| HXD (1, 6-Hexanediol) | Sigma-Aldrich, MO |
| HydraGreen™ Safe DNA Dye | ACTgene, Piscataway, NJ |
| Hyper Broth™ | Athena Enzyme Systems™, Baltimore, MD |
| IPTG (Isopropyl thiogalactosidose) | Fisher, Fair lawn, NJ |
| L- glutathione reduced | Sigma-Aldrich, MO |
| Lysozyme | Bioshop, ON |
| MPD (Methyl-pentane-diol) | Hampton Research, Aliso Viejo, CA |
| pGEX-6P-3 expression vector | GE healthcare Life Science, PA, USA |
| Poly-ethylene glycol (PEG) Molecular weight 400, 3500,6000, 8000 | Hampton Research, Aliso Viejo, CA |
| Prescision Protease | Amersham Biosciences, NJ |
| TCEP (Tris(2-carboxyethyl) phosphine) | Hampton Research, California |

| | |
|---|---|
| Wizard I & II | Rigaku, Bainbridge Island, WA |
| PACT suit | Qiagen, Toronto, ON |
| JCSG + suit | Qiagen, Toronto, ON |
| AmSO4 + suit | Qiagen, Toronto, ON |
| PEG/ION (HR2-126) | Hampton Research, Aliso Viejo, CA |
| 24- well crystal plates (HR3-170) | Hampton Research, Aliso Viejo, CA |
| HR2-110 Crystal Screen Kit | Hampton Research, Aliso Viejo, CA |
| QIAprep spin plasmid mini prep kit (27104) | Qiagen, Toronto, ON |
| QIAquick Gel extraction kit (28704) | Qiagen, Toronto, ON |
| Plasmid Maxi prep kit | Qiagen, Toronto, ON |
| QuikChange® Site-Directed Mutagenesis Kit | Agilent technologies |
| FPLC | Amersham Pharmacia biotech, Ramsey, Minnesota |
| Centrifuge J2-HS | Beckman Coulter, Mississauga, ON |
| JLA 16.250, JA 20, JLA 8.1 | |
| Gryphon | Art Robins Instruments, Sunnyvale, CA |
| Cell disruptor | Constant cell disruption systems, Daventry, UK |
| PCR MyCycler Thermal cycler | BioRad, Hercules, CA |
| Sorvall Biofuge primoR | Thermo Scientific™, Waltham, Massachusetts, USA |
| Microfuge 18 centrifuge | Beckman Coulter, Mississauga, ON |
| Centrifugal filter units | Merck Millipore, Ireland |
| Nanodrop spectrophotometer | Thermo Scientific Nanodrop 2000c |
| Vacuum concentrator | Labconco, Kansas City, MO |
| Gel imager | Syngene Imager, Frederick, MD |

# List of Abbreviations

Amp – Ampicillin

AMP-PNP – Adenylyl imidodiphosphate

BLAST – Basic Local Alignment Search Tool

CAPS – N-cyclohexyl-3-aminopropanesulfonic acid

cryo-EM – Cryo-electron microscopy

CV- Cell volume

DTT – dithiothreitol

EDTA – ethylenediaminetetraacetic acid

FPLC – Fast Performance Liquid Chromatography

GS – Glutathione Sepharose

GST – Glutathione-S-transferase

HEPES – 4-(2-hydroxyethyl)-1-piperazine ethanesulphonic acid

HXD – 1, 6-Hexanediol

IPTG – isopropyl-β-D-thiogalactopyranoside

LB – Luria-Bertani broth

MALS - Multiangle Light Scattering

MALT - mucosa-associated lymphoid tissue

MCS – multiple cloning site

MPD – 2-methyl-2, 4-pentanediol

$OD_{600}$ – optical density measured at 600 nm

PBS – phosphate buffered saline

PCR – polymerase chain reaction

PDB – Protein Data Bank

PEG – Polyethylene Glycol

PMF - Proton Motive Force

PMSF – phenylmethylsulfonyl fluoride

SDS-PAGE – sodium dodecyl sulphate polyacrylamide gel electrophoresis

SEC – size exclusion chromatography

SQ – Source Q Anion exchange column

T3SS – Type III secretion systems

TEMED – N,N,N’,N’-Tetramethylethylelediamine

Tris-HCl – tris(hydroxymethyl)aminomethane hydrochloride

w/v – weight per volume

# List of Figures

# List of tables

# 1. Introduction

## 1.1 Characteristics of the human gastric pathogen, *Helicobacter pylori*

The idea that the human stomach is a sterile organ, was changed with the discovery of *Helicobacter pylori* from gastric biopsies by Warren and Marshall in 1983 (Warren and Marshall, 1983). Initially, this bacterium was named as *Campylobacter pyloridis* but in 1989, after comparing the DNA sequence with other data, it was shown that this bacterium does not belong to the *Campylobacter* genus. It was then placed in its own genus, *Helicobacter*. The name pylorus means pyloric valve (the valve opening to the duodenum) (Mobley *et al.*, 2001).

### 1.1.1 Classification, gene organization and general metabolism in *H. pylori*

The human gastric pathogen, *Helicobacter pylori* is a Gram-negative member of the class Epsilon proteobacteria. It infects the gastric mucosa of more than 50% of the world's population thus it is the most prevalent of all the bacterial infections. In developing countries, the prevalence can be as high as 80-90%, whereas, in industrialized countries, it is lower, ranging between 10-50 % (Rothebacher and Brenner, 2003). It causes chronic inflammation of the inner wall of the stomach i.e. gastritis and it is the main cause of peptic ulcers (Blaser, 2006). *H. pylori* is also associated with stomach cancer and a rare type of cancer called MALT (mucosa-associated lymphoid tissue) lymphoma. In cancer-related deaths worldwide, gastric adenocarcinoma is the second leading cause (Ferlay *et al.*, 2010).

The International Agency for Research into Cancer (IARC) (a part of the World Health Organization - WHO) has classified *H. pylori* as a "Class I Carcinogen". Along with gastric diseases, studies also suggest that *H. pylori* might also be a causative agent of some non-gastric diseases such as chronic liver disease and vascular disease (Pellicano *et al.*, 2009). *H. pylori* have evolved strategies to downregulate the activation of host immune cells and escape immune effector cells (Blaser and Antherton, 2004). *H. pylori* is motile by means of several polar sheathed flagella which provide locomotion. The flagellum is also necessary for the initial gastric colonization of the bacterium (Otemann and Lowenthal, 2002). The study of the effect of virulence on gnotobiotic piglets strongly correlates motility to the virulence of some bacteria like *C. pylori* (Eaton *et al.*, 1989). Therefore, the basic biology and biochemistry of flagellar protein export mechanism in flagella are of value in understanding *H. pylori* infection and gastric colonization.

### 1.1.2 Bacterial motility

The unusual motility of *H. pylori* in the viscous environment of the gastric mucosa is hypothesized to be due to the polar flagella and the helical cell shape. The flagella provide propulsive torque as well as a rotary movement while the helical cell shape creates a corkscrew-like rotation, which allows bacteria to promote motility in the viscous mucus (Berg *et al.,* 1979; Ferrero and Lee, 1988). The helical shape is the resultant of cell elongation, curvature and twist. However, study of *H. pylori* mutants with less helical cell bodies showed that it retains the capacity of high swimming velocity in viscous liquids, suggesting that helical shape may not be a critical factor for motility (Bonis *et al.,* 2010; Sycuro *et al.,* 2010). Flagellar motility is a very important determinant for colonization by *H. pylori*, presumably allowing migration towards the gastric mucus as well as allowing movement within the mucus to reach microenvironments favorable to growth. *Helicobacter* species demonstrate a high locomotion rate in viscous environments compared with many motile bacteria. *Salmonella* and *E. coli* (peritrichous, rod shaped bacteria) swim with highest velocities in low to slightly viscous substances but their velocity decreases with increase in the viscosity of the environment (Lertsethtakarn *et al.,* 2011). In contrast, *H. pylori* (polar flagella, helical bacteria) demonstrate approximately twice the velocity of rod-shaped bacteria in slightly viscous substance and can retain these velocities even in a highly viscous environment (Lertsethtakarn *et al.,* 2011).

### 1.1.3 *H. pylori* motility, colonization of gastric mucosa and virulence

Persistent colonization and infection of gastric mucosa by *H. pylori* is possible due to motility, physical environment of the stomach, biochemical factors, virulence factors, as well as genomic diversity of the bacterium. The microenvironment of the gastric lumen and mucosa contains a variety of biochemical factors which are hostile to bacteria. *H. pylori* can grow in liquid environment around pH 5.1, meaning the bacteria can survive with the acidity of stomach after being ingested orally with food. After oral ingestion, *H. pylori* must survive the acidic lumen of the stomach until penetration of the gastric mucus, after which it will have to survive in the pH gradient of the gastric mucus layer. The activity of the enzymes such as urease, formamidase and arginase which are associated with production of ammonium ions help the bacteria to endure the acidic environment. For hydrolysis of urea present in the gastric mucus layer, *H. pylori* uses urease enzyme and produces $NH_3$. This $NH_3$ rapidly diffuses to the mucus

layer and increases the pH, thus allowing penetration into the mucus layer by coupling the corkscrew like propulsive rotation originating from the helical shape and flagellar motility.

Mutant *H. pylori* lacking functional flagella showed loss of mouse infectivity demonstrating that functional flagella are required to colonize the host's stomach (Schoenhofen *et al.,* 2006). The study of the effect of virulence on gnotobiotic piglets strongly correlates motility to the virulence of *H. pylori* (Eaton *et al.,* 1989). Adherence of *H. pylori* to the gastric mucosa through bacterial envelope structures like adhesins also plays a critical role in initial colonization. Diversification including intragenomic and intergenomic changes of the *H. pylori* genome is likely one of the mechanisms for immunological evasion and adaptation to the gastric environment (Rhee *et al.,* 2014). Two major pathological proteins *viz*. CagA (cytotoxin-associated gene A) and VacA (vacuolating cytotoxin A) are determining factors in pathogenesis due to *H. pylori.* VacA binds to gastric epithelial cells and is internalized, causing cellular vacuolation characterized by the accumulation of large vesicles which have the characteristics of both late endosomes and early lysosomes (Yamaoka, 2010; Palframan *et al.,* 2012). CagA when injected into the cytosol of the host gastric epithelial cells, triggers signal transduction (e.g. proliferation and inflammation), leading ultimately to gastric disease (Yamaoka *et al.,* 2010; Lai *et al.,* 2006). In conclusion, a number of host and bacterial factors are responsible for colonization and persistence of *H. pylori* in the human stomach ultimately resulting in pathogenesis and clinical symptoms.

## 1.2 The Type III Flagellum Export System

### 1.2.1 Type III secretion systems

Many bacteria have evolved specialized nanomachines that have the capacity to modulate a variety of cellular activities and are used to deliver effector proteins into eukaryotic host cells (Alvarez *et al.,* 2009; Galàn, 2009; Llosa *et al.*, 2009) (Figure 1.1). Macromolecular nanomachines help the bacteria to perform complex biological tasks like locomotion and protein translocation. One such nanomachine present in the Gram-negative bacteria is type III secretion system (T3SS). It is a specialized protein-export apparatus that helps bacteria to transport proteins across the cell membrane and extracellular spaces. T3SSs are the best characterized amongst the protein injection machines (Buṫtner, 2012; Cornelis, 2006; Galàn and Wolf-Watz, 2006). T3SSs are widespread in nature and play vital roles not only for transport of effector

proteins but also in symbiotic interactions with plant, vertebrate or insect hosts (Buẗtner and He, 2009; Correa *et al.*, 2012; Silver *et al.*, 2007). Although this secretion machine is highly conserved across most Gram negative bacterial species, the effector proteins that are delivered by these nanomachines are specific for each pathogen that encodes them (Galàn, 2009).

The T3SS can be broadly classified into two classes: the flagellar T3SS which promote cell motility (Erhardt *et al.*, 2010) and the non-flagellar T3SS (NF-T3SS) that promote bacterial pathogenesis by injecting effector proteins directly into eukaryotic host cells (Abby and Rocha, 2012; Buẗtner, 2012; Erhardt *et al.*, 2010). We will discuss both types of T3SS in the following sections. The core of the T3SS is the needle complex which is composed of a multi-ring base substructure and a needle-shaped extension (Chatterjee *et al.*, 2013; Kosarewicz *et al.*, 2012; Loquet *et al.*, 2012; Marlovits *et al.*, 2006; Marlovits *et al.*, 2004; Radics *et al.*, 2014; Worral *et al.*, 2011). The base is approximately 25 X 30 nm and is composed of two inner membrane rings (IR1 and IR2) connected to two outer membrane rings (OR1 and OR2) through the neck (Figure 1.2) (Galàn *et al.*, 2014).

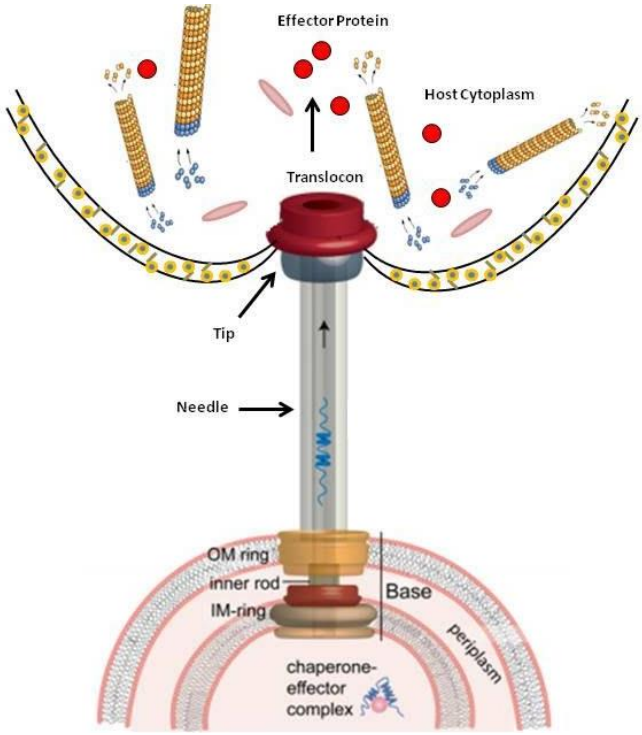


Figure 1.1. **Cartoon representation of an injectisome T3SS.**

The T3SS showing the needle complex delivering the effector proteins into the host cell. Figure adapted from Chatterjee *et al*., 2013.

The lower rings are composed of two proteins (YscJ/PrgK/MxiJ and YscD/PrgH/MxiG in *Yersinia*, *Salmonella* and *Shigella* spp., respectively) (Table 1.1), whereas the outer ring and neck are composed of just one protein (YscC/ InvG /MxiD) (Galàn *et al.*, 2014).

The needle is approximately 50 nm long and it is connected to the base through the inner rod. The inner rod docks into a socket-like structure within the base. The complete structure is traversed by a channel which is approximately 20 Å in diameter that transports proteins (Radics *et al.*, 2014). The needle is assembled from approximately 100 copies of a single subunit, PrgI/YscF/MxiH (80 subunits) which are arranged in a helical fashion (Cordes *et al.*, 2003; Galkin *et al.*, 2010; Kubori *et al.*, 2000; Loquet *et al.*, 2012) The tip complex is presumably involved in the sensing of the target cells as well as in the distribution of the protein translocases (Epier *et al.*, 2012; Mueller *et al.*, 2005). The proximal end is connected to an export apparatus (EA) as well as to an ATPase complex in the cytosol. The export apparatus consists of a group of proteins located at the center of the inner rings that serves as a channel for mediating the passage of type III secreted proteins through the inner membrane (Wagner *et al.*, 2010).

All T3SSs show the presence of five highly conserved inner membrane proteins (Table 1.1) (InvA, SpaP, SpaQ, SpaR, and SpaS in *Salmonella* spp.) (flagellum counter parts- FlhA, FliP, FliQ, FliR, FlhB respectively) which are essential for the function of T3SS (Allaoui *et al.*, 1994; Fields *et al.*, 1994; Galàn *et al.*, 1992; Ginocchio and Galàn, 1995; Groisman and Ochman, 1993; Plano *et al.*, 1991). Finally, many cytoplasmic proteins are attached to the cytoplasmic side of the needle complex and are essential for secretion. Many of these proteins are well conserved among all T3SSs (Collazo *et al.*, 1996; Collazo *et al.*, 1995; Eichelberg *et al.*, 1994; Fields *et al.*, 1994; Morita *et al.*, 2006; Woestyn *et al.*, 1994). As recently suggested by electron tomographic studies of the needle complex *in situ*, these proteins are presumed to be forming a defined structure, (Abrusci *et al.*, 2013; Kawamoto *et al.*, 2013; Kudryashev *et al.*, 2013). An ATPase complex is another highly conserved cytosolic component. Finally, the passage of effector proteins into eukaryotic cells is mediated by the protein translocases (SipB/SipC, in *Salmonella*) (Collazo and Galán, 1997; Rosqvist *et al.*, 1995). Although so much information is available about the structure of the T3SS, there are some protein densities observed in the high-resolution cryo-EM map that are still not accounted for (Galàn *et al.*, 2014).

Some of the protein densities are most likely to represent transmembrane domains of components of the export apparatus, whose crystal structures are not yet available. Remarkable progress in the last ten years made it possible to unravel part of the mystery of the structure and organization of the type III secretion machine. Finally, the importance of the T3SS in bacterial pathogenesis has stimulated efforts to design/develop novel therapeutic strategies. It is speculated that during the next few years these efforts will begin to translate into effective therapeutics.

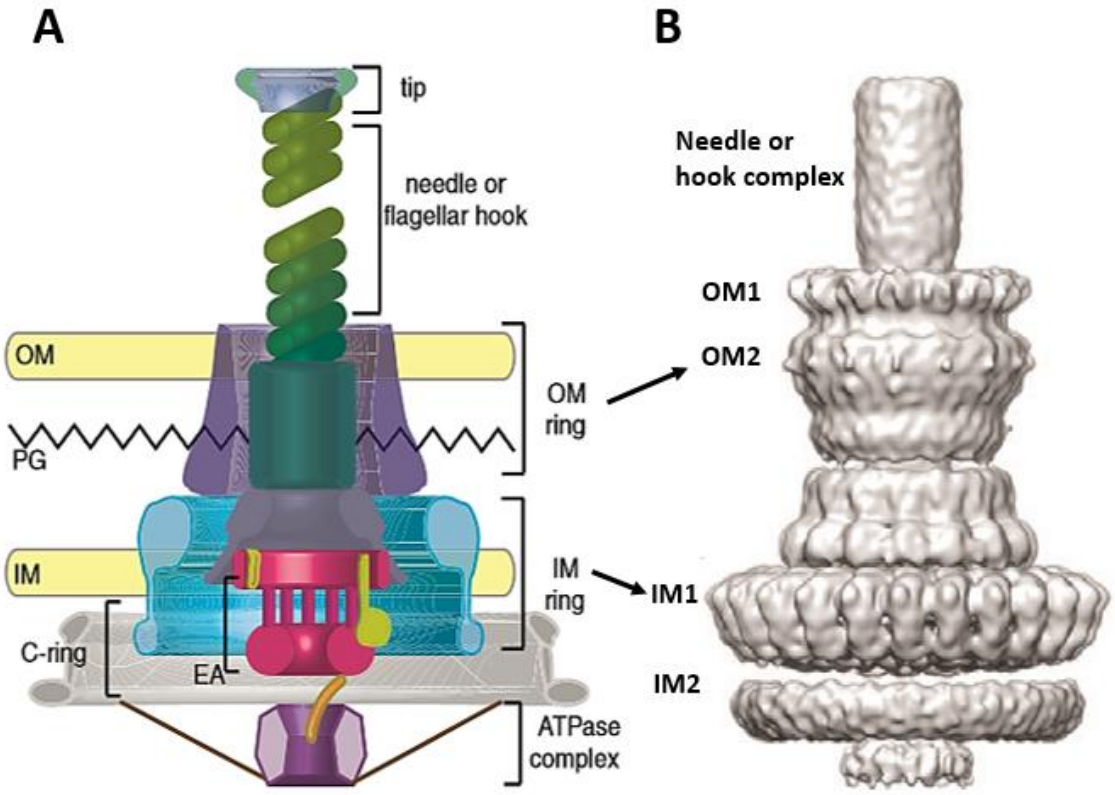


Figure 1.2. **General schematic diagram and cryo-EM image of the *Salmonella* T3SS.**

A. The cartoon representation of a typical T3SS showing major substructure (the export apparatus (EA), the needle/hook, the tip and the ATPase complex) localization with respect to the outer membrane (OM), inner membrane (IM) and peptidoglycan layer (PG). B. Surface image of the 3-D reconstruction of the cryo–electron microscopy (cryo-EM) map of the needle complex from *S. typhimurium* showing different substructures. Figure adapted from (Galàn *et al*., 2014)

**Table 1.1. Nomenclature of specific T3SS secretion components from commonly studied bacteria.**

A previously suggested Sct nomenclature system is given for reference. The table is adapted from Abrusci *et al*., 2014.

| ***Sct* common nomenclature** | ***Yersinia spp.*** | ***Shigella spp.*** | ***Salmonella* SP-1** | **Flagellar apparatus** | **Proposed function** |
|---|---|---|---|---|---|
| **SctF** | YscF | MxiH | PrgI | - | Needle filament component |
| **SctI** | YscI | MxiI | PrgJ | - | Inner rod protein |
| **SctC** | YscC | MxiD | InvG | - | Needle complex outer rings |
| **SctD** | YscD | MxiG | PrgH | - | Needle complex inner rings |
| **SctJ** | YscJ | MxiJ | PrgK | FliF | Needle complex inner rings |
| **SctR** | YscR | Spa24 | SpaP | FliP | Export apparatus |
| **SctS** | YscS | Spa9 | SpaQ | FliQ | Export apparatus |
| **SctT** | YscT | Spa29 | SpaR | FliR | Export apparatus |
| **SctU** | YscU | Spa40 | SpaS | FlhB | Export apparatus; substrate switching |
| **SctV** | YscV | MxiA | InvA | FlhA | Export apparatus |
| **SctK** | YscK | MxiK | OrgA | FliG | Cytoplasmic sorting platform |
| **SctQ** | YscQ | Spa33 | SpaO | FliM+ FliN | Cytoplasmic sorting platform |
| **SctL** | YscL | MxiN | OrgB | FliH | Links ATPase to sorting platform |
| **SctN** | YscN | Spa47 | InvC | FliI | ATPase |
| **SctO** | YscO | Spa13 | InvI | FliJ | Flagellum specific Chaperone |
| **Sct-P** | YscP | Spa32 | InvJ | FliK | Inner rod assembly; substrate switching |

### 1.2.2 Structure of the bacterial flagellum and its T3SS

The majority of the swimming bacteria use a reversible, self-assembling, rotary motor called the flagellum (Figure 1.3). Bacteria can actively move using this nanomachine. There are a number of different proposed models for assembly, structure and function of this rotary motor (Soutourina and Bertin, 2003; Pallen and Matzke, 2006). In enteric bacteria flagellar assembly is a very tightly regulated process (Chevance and Hughes, 2008; Macnab, 2003; Aldridge and Hughes, 2002; Chilcott and Hughes, 2000).

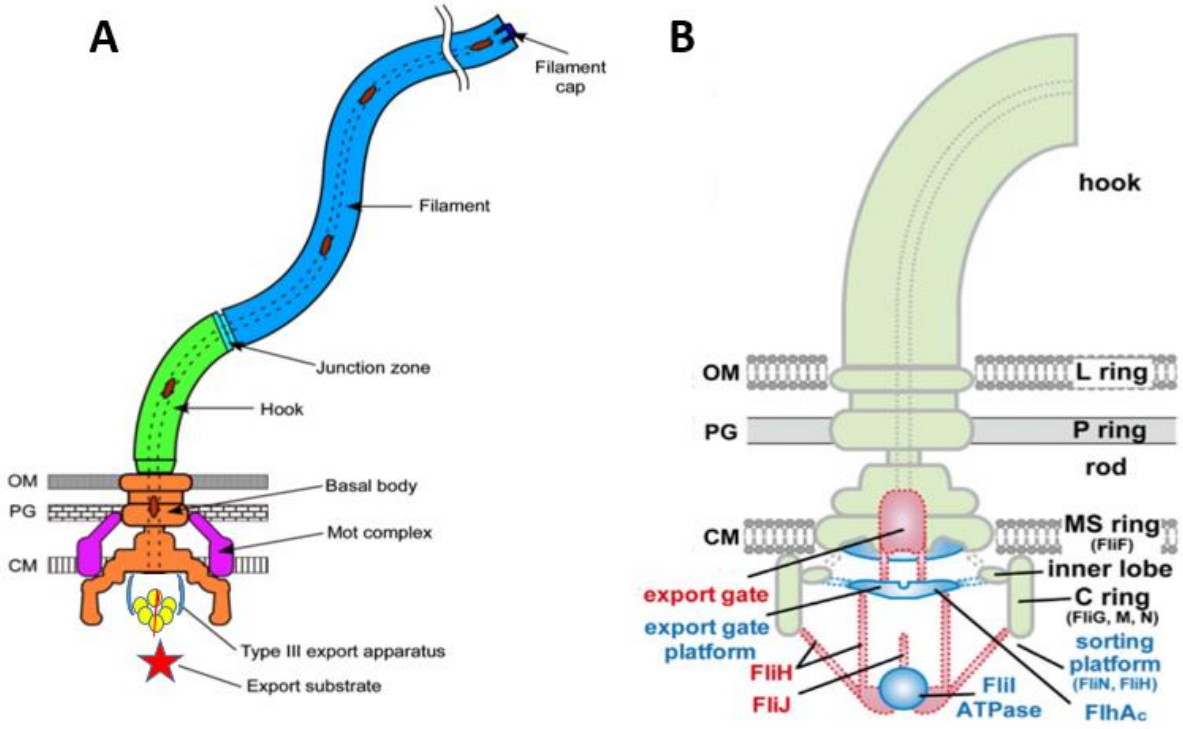


Figure 1.3. **Schematic diagram of the flagellum.**

**A**. The flagellum consists of the basal body, the hook, the hook-filament junction zone, the filament and the filament cap. Most flagellar component proteins are transported using a specific export apparatus from the cytoplasm to the distal end of the growing flagellar structure where their assembly occurs. OM, outer membrane; PG, peptidoglycan layer; CM, cytoplasmic membrane (Figure from Minamino, 2014). **B**. Components of the flagellum and flagellar export apparatus showing their specific positions on the flagellum. Figure from (Kawamoto *et al*., 2013).

*E. coli* and *Salmonella* (in Gram-negative bacteria) and *Bacillus* (Gram-positive) are the major model organisms that have provided detailed insights into understanding the structure, assembly, and function of the bacterial flagellum (Minamino and Imada, 2015). Flagella respond to external stimuli by a mechanism called chemotaxis. (Bourret and Stock, 2002; Bren and Eisenbach, 2000). The flagellum is mainly divided into three parts: the basal body rod, the hook and the filament (Figure 1.3). The basal body contains the flagellar motor and also includes a rotor, a stator, a rod and a bushing complex surrounding the distal rod. The flagellar motor contains approximately 40 different types of proteins. The motor generates torque from ion flux across the inner membrane thus facilitating flagellum rotation (Chen *et al.*, 2011). The rotor is present in the inner cell membrane, and it is attached to the flagellar filament via the hook (universal joint). The filament is made up of about 20,000 copies of the flagellin proteins which are exported to the tip of the filament through the central channel passing through an annular pore located in the basal body (Fan *et al.*, 1997). These flagellins collectively form a helical structure with a hollow core, approximately 15 nm in diameter and 15µm in length (Berg *et al.*, 2002). The scaffold present at the tip is made up of five copies of the hook-associated proteins that help in folding and polymerization of flagellin subunits (Merz *et al.*, 2000).

The flagellar basal body consists of the MS ring, rod, and L- and P- rings. An outer ring known as the L-ring (lipopolysaccharide) is present only in gram-negative bacteria. The P ring is present in the peptidoglycan layer. Due to the absence of an outer membrane, motile Gram positive bacteria do not possess the outer pair of rings. The flagellar basal body consists of a specialized protein export apparatus that is required for assembly of the flagellum. It is a highly dynamic structure coordinating protein export. The export apparatus is composed of a transmembrane export gate (made up of FlhA, FlhB, FliO, FliP, FliQ, FliR) and a cytoplasmic ATPase complex composed of FliH, FliI and FliJ (Minamino *et al.*, 2008a; Minamino, 2014). We will discuss the detailed structure of the export apparatus in the next section.

The flagellar rotary motor is powered by a proton gradient across the cytoplasmic membrane (Meister *et al.*, 1987; Manson *et al.*, 1977). Certain alkalophilic organisms, as well as some bacteria like *Bacillus*, utilize a $Na^+$ driven motor (Imae and Atsumi, 1989). Export of flagellum proteins requires both ATP and the membrane Proton Motive Force (PMF) (Macnab,

1996; Paul *et al.*, 2008). For flagellar secretion and rotation, PMF is required whereas ATP hydrolysis might be limited to substrate delivery.

### 1.2.3 Flagellar T3SSs

The secretion of extracellular components of the flagellum is carried out by their own specialized T3SS. Flagellar export in *Salmonella*, *E. coli*, *B. subtilis*, *H. pylori* and other bacteria uses a specialized T3SS present at the base of the flagellum. There are sequence and structural similarities between bacterial flagella, T3SS and their subunits demonstrating that the flagellum T3SS and the Gram negative pathogenic T3SSs very likely derived from a common ancestor (Blocker *et al.*, 2003).

Secretion of flagellar proteins requires at least 11 components of the export apparatus (Figure 1.4). Out of 11 components, 7 are membrane proteins (FlhA, FlhB, FliF, FliO, FliP, FliQ and FliR) that are present in the flagellar basal body (Fan *et al.*, 1997; Minamino and Macnab,1999). Three other proteins (FliI, FliH and FliJ) are soluble cytoplasmic components, forming the ATPase complex. FliI ATPase is thought to provide energy for export by converting energy of ATP hydrolysis into energy for export of proteins (Fan *et al.*, 1997). FliH is a negative regulator of FliI; and FliJ is a general flagellar secretion co-chaperone (Minamino and Macnab, 2000b). Remarkably, FliH shows sequence similarity to the b & δ subunits of $F_OF_1$ ATPase, while FliI shows significant sequence similarity to the catalytic β- subunit of $F_OF_1$ ATPase (Albertini *et al.*, 1991; Vogler *et al.*, 1991; Lane *et al.*, 2006), including some well-conserved residues present in the Walker box motifs (Minamino and Macnab, 2000b; Imada *et al*., 2007). In the case of *H. pylori* and other Epsilon proteobacteria there is little evidence of the presence of FliJ. However, CJJ81176_1489 from *Campylobacter jejuni* showed low but significant sequence similarity to FliJ hence, it was proposed that CJJ81176_1489 is a homolog of FliJ (Gao *et al.*, 2014). Deletion of this gene in *Campylobacter jejuni* resulted in complete absence of flagella suggesting that this gene is essential for flagellar assembly. A recent bioinformatics and motility analysis have identified HP0256 as a potential FliJ homologue (Douillard *et al*., 2010) and is required for full motility.

There are number of differences between *Salmonella* and *H. pylori* flagellar export systems. In *H. pylori* the flagellar genes are under the control of independent promoters. The genetic organization of 40 genes which are required for flagellar assembly and/or function is

completely different in *H. pylori* as compared to other bacteria. Many flagellar genes appeared to be scattered in *H. pylori* as opposed to the well-defined (clustered) organization in *Salmonella, E. coli etc*. Furthermore *H. pylori*, has a number of atypical features in the flagellar genetics which makes *H. pylori* a unique system. We will discuss *H. pylori* specific flagellar system in section 1.3.

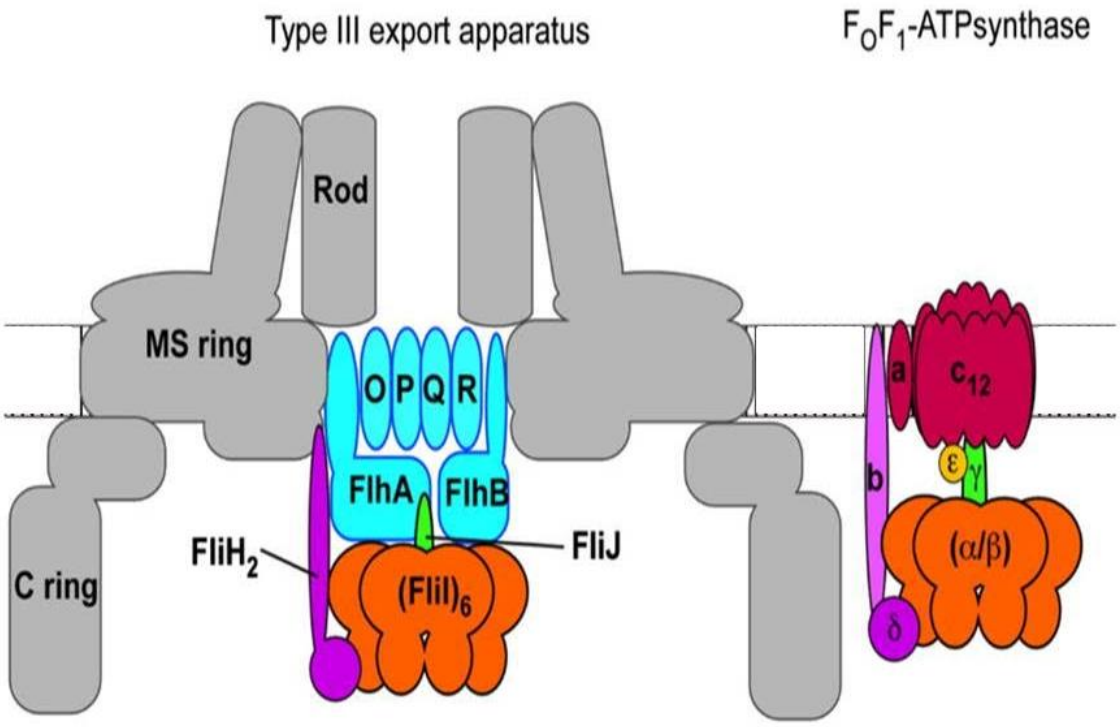


Figure 1.4. **Cartoon representation of a comparative model of the flagellar type III export apparatus and $F_OF_1$ ATP synthase.**

Transmembrane domain of the export apparatus is composed of FlhA, FlhB, FliO, FliP, FliQ and FliR that are located within MS ring. $FliI_6$-FliJ ring complex shows remarkable similarity to the $\alpha_3\beta_3\gamma$ ring complex of $F_OF_1$-ATPsynthase. FlhA anf FlhB forms the export gate. FliH anchors the $FliI_6$-FliJ ring complex to the export gate. FliH is homologous to b and δ subunits of peripheral stalk of $F_OF_1$-ATP synthase. (Figure from Minamino, 2014).

### 1.2.4 Flagellum and injectisome T3SS: Morphological similarities and shared genes

The evolutionary relation between the flagellum and the Type III injectisome has been always a subject of debate. Early phylogenetic studies suggested that both these systems share a common ancestor while recent studies indicate that modern injectisomes are derived from non-flagellar ancestors (Diepold and Armitrage, 2015). The flagellum and the injectisome are debatabely the most fascinating and complex nanomachines found in bacteria. While the term 'T3SS' is sometimes used to represent the whole injectisome, in this study it will be used to represent the export machinery present in both, 'the flagellum' as well as 'the injectisome'.

Similar to the injectisome, the flagellar basal body (FBB) is made up of a multi-ring structure that spans both the bacterial membranes and the peptidoglycan (Minamino and Namba, 2008) (Figure 1.5). A part of the T3SS machinery, the injectisome resembles Hook-Basal body (HBB) complex of the flagellum (Kubori *et al.,* 1998; Thomas *et al*., 2001). The bacterial injectisome and flagellum both consists of large assemblies of 20-30 different proteins. These proteins are exported to the distal ends in both nanomachines with the help of a specialized, conserved protein export machinery classified as T3SS. Their basal bodies house the T3SS export apparatus (Hueck, 1998; Minamino and Namba, 2008). Injectisome and FBB isolated from bacterial cells share a highly similar architecture as viewed by electron microscopy (Kubori *et al*., 1998; Schraidt *et al*., 2011; Thomas *et al*., 2006). Also, many of their component proteins are homologous (Hueck, 1998; Kosarewicz *et al*., 2012) which strongly suggests that injectisomes are evolutionally related to the flagella. Both T3SS shows the presence of a membrane-embedded export gate and a cytoplasmic protein complex with an ATPase.

Table 1.1 shows the common proteins present in both the systems. In the injectisome, five membrane proteins make up the export gate. The proteins SctR, SctS, SctT, SctU and SctV, form the export gate in the injectisome while the export gate of the *Salmonella* flagellar protein export apparatus is made up of six membrane proteins, namely, FlhA, FlhB, FliO, FliP, FliQ, FliR. The InvC ATPase (homolog of SctN ATPase) forms a homo-hexameric ring on the cytoplasmic side with OrgA, OrgB and InvI (homologs-SctK, SctL, SctO respectively) (Kawamoto *et al*., 2013; Hu *et al*., 2015). The FliI from flagellum export apparatus forms a hexamer ring complex with FliH and FliJ (Minamino *et al*., 2008a). In injectisome, the export gate components are present in the central part of the inner membrane ring (IR) complexes

(Wagner *et al*., 2010). SctQ forms a sorting platform by providing specific binding sites for SctK and SctL. SctN ATPase is important for chaperone release and substrate unfolding. The C-ring (composed of FliG, FliM and FliN) is clearly visible in the flagellum system as a part of the switch complex. In case of injectisome, though there are homologous proteins, no C-ring structure was detected (Kawamoto *et al*., 2013; Kudryashev *et al*., 2013). Recent genetic studies and evolutionary analysis have suggested that injectisome evolved from the flagellum (Abby and Rocha, 2012) and then adapted to different eukaryotic cells but the core structure was conserved throughout the process.

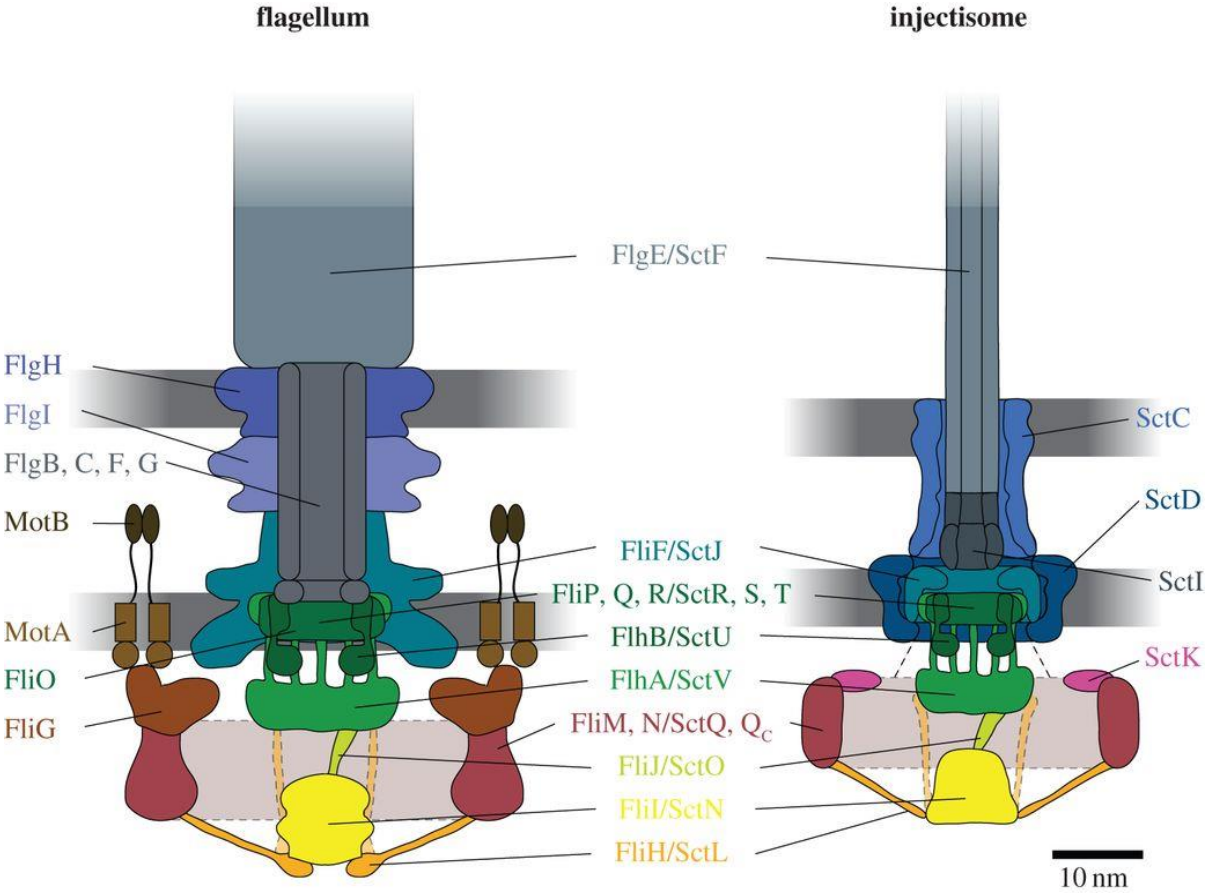


Figure 1.5. **Structural model of the flagellar as well as injectisome T3SS.**

Structural components listed in the center are present in both T3SS. Flagellum specific components are listed on the left side whereas injectisome specific components are listed on the right side. Figure from (Diepold and Armitage, 2015).

### 1.2.5 Common evolutionary relation between rotary ATPases and the flagellar export apparatus

F-V family ATPases/ATP synthases couple ATP hydrolysis or synthesis to ion/proton translocation across the cellular membranes. Rotary ATPases are found across all life forms and the overall architecture in all of them is similar. On functional properties and taxonomic origin, three evolutionarily subtypes of Rotary ATPase are evidenet as F, V and A type (Cross and Muller, 2004) (Figure 1.6).

Prokaryotes predominantly have either F-ATPases ($F_OF_1$) which synthesizes ATP or V-ATPases ($V_OV_1$), which functions to pump protons across the membrane. Eukaryotes have both F-ATPases and V-ATPases, which can function either to pump protons in or outside the membrane, depending on the assigned function (Nakano *et al*., 2008 and Feniouk *et al*., 2007). F-type ATP synthases are present in mitochondria or chloroplasts where they function as biological power converters (Iino and Noji, 2013). Eukaryotic V-ATPases are driven by ATP hydrolysis and mainly acidify organelles such as endosomes, lysosomes, Golgi-derived vesicles and the vacuoles (Forgac, 2007; Li and Kane 2009). As compared to prokaryotes, eukaryotes work oppositely by hydrolyzing ATP and using the energy to drive protons across the membrane (Jefferies *et al*., 2008). A-type ATPase complexes are very common in well-known archaebacteria, which are evolutionarily closer to V-ATPases and mainly hydrolyze ATP (Cross *et al*., 2004).

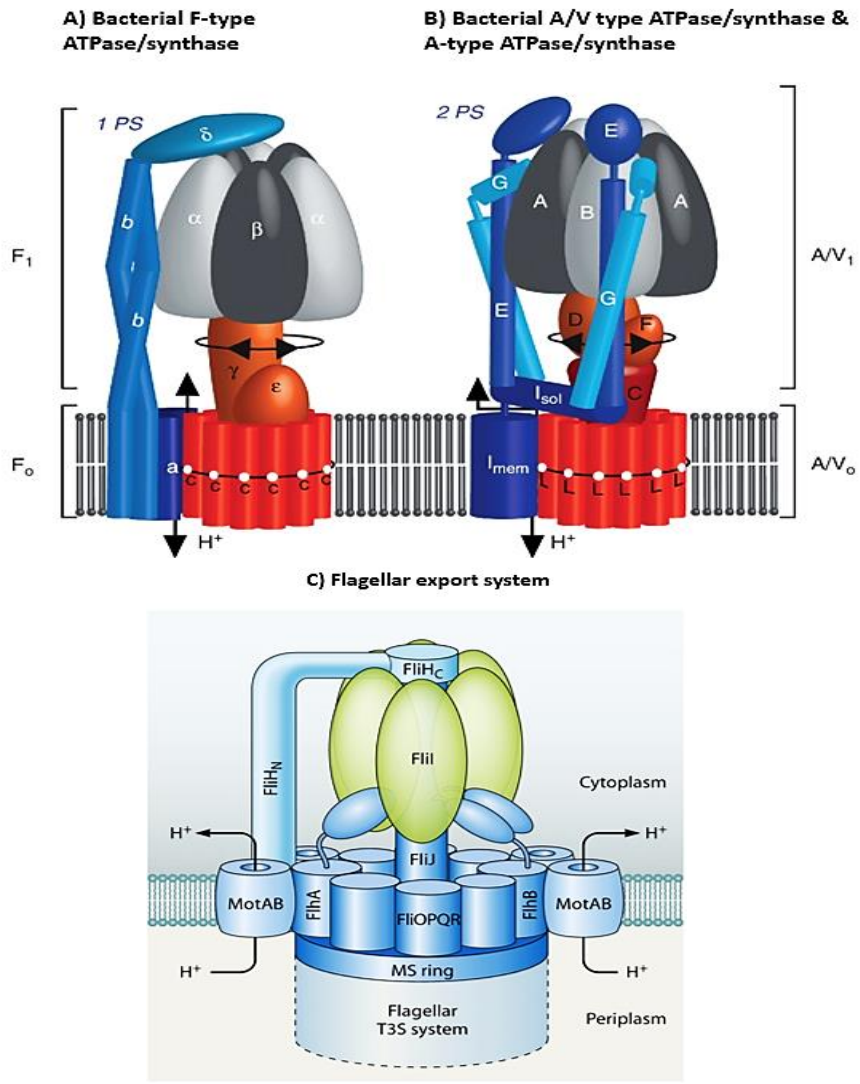


Figure 1.6**. Schematic diagrams of rotary ATPase, Flagellar export system and their major subunits.**

A. Bacterial F-type ATPases/synthases showing membrane embedded $F_o$ motor, soluble $F_1$ motor (is made up of $\alpha_3\beta_3\gamma$ subunits), and one peripheral stalk (made up of b and δ subunits). These ATPases can synthesize ATP as well as act as proton pumps when ATP levels are high. B. Bacterial V-type & archaeal A-type ATPases share a common architecture. The minimal rotary domain is comprised of $A_3B_3D$ subunits. They contain two peripheral stalks and are believed to be bifunctional. Subunits that bind to nucleotide are shown in grey color, remaining stator subunits in blue, rotary subunits in red color and protons indicated as white spheres (Figure from Stewart *et al*., 2014). C. The flagellar type III export system consists of an export gate made (FlhA, FlhB, FliO, FliP, FliQ and FliR) and a water-soluble ATPase complex (FliH, FliI, and FliJ). The $FliI_6$-FliJ structure from cryo-EM studies shows striking structural similarity to $F_1/V_1$ subunits. FliH also shows sequence similarity to peripheral stalk subunit E of V-ATPase (Figure from Büttner, 2012).

All types of rotary ATPases have two motors, one of them contains a nucleotide binding site for ATP whereas the other is a membrane bound motor for translocating protons/ions (Stewart *et al*., 2014). In F-type ATPase, the minimal ATP-driven rotor is $F_1$, which is made up of $\alpha_3\beta_3\gamma$ subunits. The γ subunit is comprised of two domains, a coiled-coil domain penetrating the central $\alpha_3\beta_3$ cylinder and another one is a globular domain with a α/β fold which is in contact with the $F_1$-ε subunit (Kishikawa *et al*., 2013). Similarly, in V-type ATPases, the minimal rotary domain is comprised of $A_3B_3D$ subunits. It has been suggested that due to the presence of long α-helices, the D-subunit is most likely to be an analog of $F_1$ γ even though there is no significant sequence similarity between them (Mulkidjanian *et al*., 2007; Kishikawa *et al*., 2013). The central stalk of all the ATPases is important for rotary catalysis, and therefore the distinct lack of conservation in the F - and V - type ATPase stalks may have a connection with their evolutionary origin. It was proposed that ancestors of these enzymes lacked a central stalk. The major difference between all the ATPases exists in the peripheral stalks. The peripheral stalk connects the cytoplasmic ATPase domain ($F_1$/ $V_1$/$A_1$) to the membrane embedded domain ($F_O$/ $V_O$/$A_O$). F-ATP-synthases contains one peripheral stalk (subunits- b and δ) (Dunn *et al*., 2000; Dickson *et al*., 2006) (Figure 1.6) while A-ATPases contain two peripheral stalks (subunits E and G) (Lee *et al*., 2010). Eukaryotic V-ATPases contains three peripheral stalks (Kitagawa *et al*., 2008). Each peripheral stalk is made up of a heterodimer of E and G subunits. Out of the three stalks, two connect the ATPase domain to membrane bound domain while third stalk is attached to C-subunit (absent in F-ATP synthases/A-ATPases) (Oot *et al*., 2012).

The FliI ATPase from the flagellar export apparatus shows extensive sequence and structural similarity to the $F_1$- α/β and $V_1$-A/B subunits (Imada *et al*., 2007). FliJ shows a striking structural similarity to the $F_1$-γ subunit (Ibuki *et al*., 2013). The coiled-coil domain of FliJ is also present in the non-flagellar Type-III homolog YscO-like proteins (Lorenzini *et al*., 2010). FliJ promotes $FliI_6$ ring formation by binding to the center of the hexameric ring thus facilitating FliI ATPase activity (Minamino *et al*., 2008b). Similar arrangement is found in the $F_OF_1$-ATPase where the antiparallel α helical coiled-coil of the γ subunit penetrates into the center of the $\alpha_3\beta_3$ ring. Kishikawa *et al*., have proposed that FliJ, $V_1$-D, and $F_1$- γ all are evolved from a common evolutionary origin (Kishikawa *et al*., 2013).

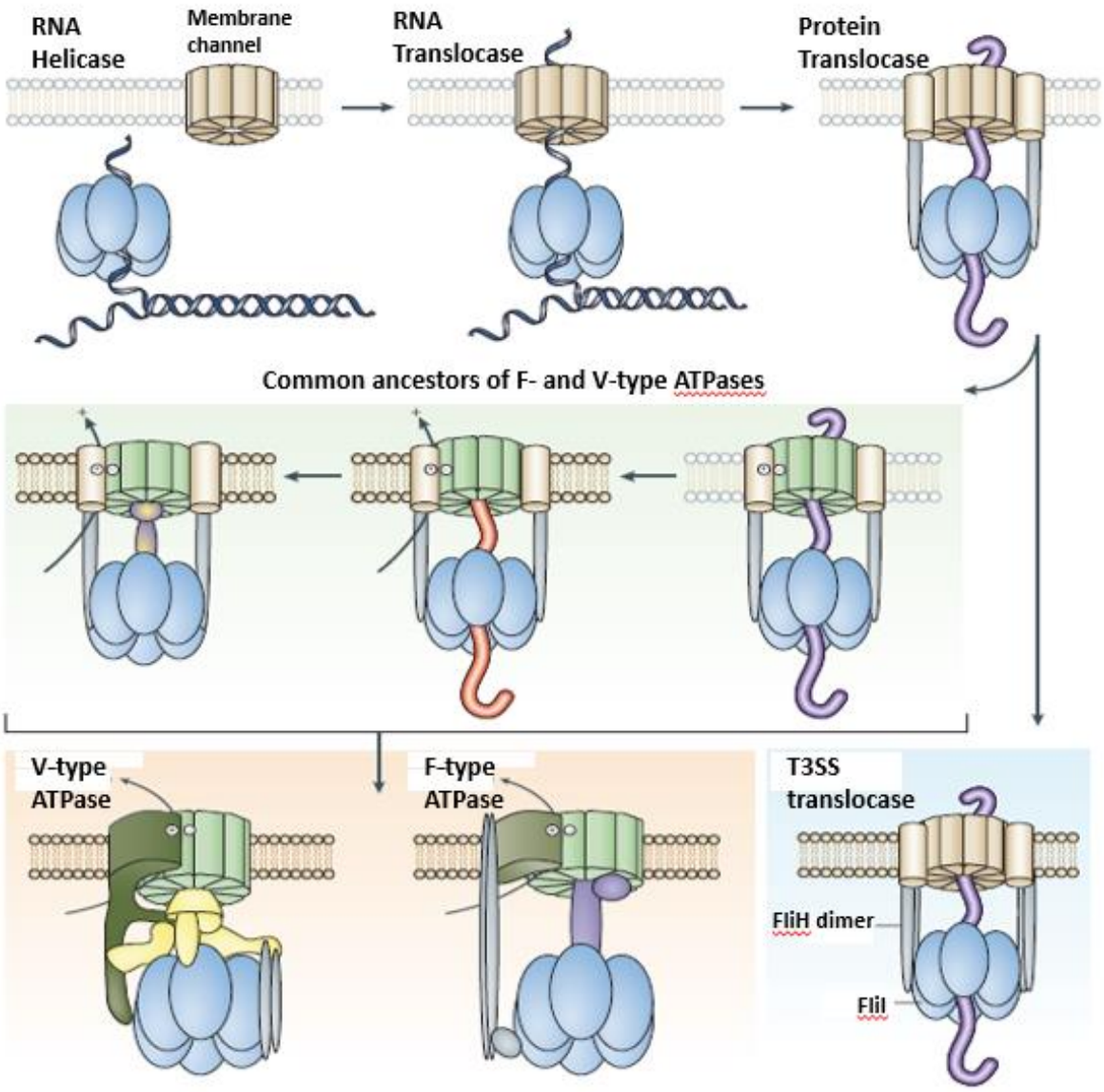


Figure 1.7**. Proposed evolutionary relation between F- and V-type ATPases and flagellar T3SS.**

Ancestral, hypothetical semi-permeable membranes are shown by empty circles whereas modern, ion-tight membranes are shown in brown circles and dense curved lines. The proposed hypothesis regarding two peripheral stalks in the primordial protein translocase as well as in the T3SSs is based on the assumption that only one peripheral stalk (in the absence of the translocated substrate) would make the translocase unstable. The central panel shows the common ancestors of F- and V-type ATPases which are the three consecutive evolutionary intermediates. The translocated, partially unfolded protein is denoted by the purple tubes. The translocated protein that is trapped in the membrane translocase channel is shown in red colored tube. (Figure was adapted from Mulkidjanian *et al*., 2007)

Based on all the above facts it was suggested that the type III export system share similarity with rotary ATPases and share a common evolutionary origin with the bacterial flagellar and non-flagellar type III export system (Figure 1.7) (Mulkidjanian *et al*., 2007). They have proposed that the origin of the rotary cation-translocating ATPases started with an RNA helicase (Rho transcription terminator) and a membrane channel. Even though there needs to be more evidence in this field, based on all structural and sequence comparison it can be stated that rotary ATPases and flagellar export apparatus very likely share a common evolutionary origin.

## 1.3 *H. pylori* specific flagellum proteins

### 1.3.1 *H. pylori* flagellum export apparatus

Recent *in-situ* cryoelectron tomography studies in *C. jejuni* and *H. pylori* have shown that the flagella of these organisms are among the most diverse flagella since they have unique ultrastructural components (Gao *et al*., 2014). The understanding of flagellar export apparatus of *H. pylori* is still somehwhat limited as compared to that in *E. coli* and *Salmonella.* There is evidence identifying many proteins (FliI, FliH, FliP, FliQ, FlhA, FlhB) that constitute the flagellar export locus in *H. pylori* (Porwollik *et al*., 1999). Recently a FliO homolog has been identified in *H. pylori* and it is larger in size as compared to its counterparts in *E. coli* and *Salmonella* (Tsang and Hoover, 2014). The *H. pylori* FliO homolog is predicted to have a large N-terminal periplasmic domain (Tsang and Hoover, 2014). There is very little evidence for the presence of FliJ in *H. pylori* but a potential homolog HP0256 has been identified which shows 15% sequence identity with *B. subtilis* FliJ and aligns reasonably well with other FliJ homologs (Douillard *et al*., 2010). Characterizing the roles of all components of the flagellar export apparatus in *H. pylori* is crucial for understanding how this organism coordinates flagellar gene expression with flagellum assembly.

### 1.3.2 *H. pylori* flagellum gene regulation

In *H. pylori*, a complex cellular process such as flagellum formation is coordinately regulated despite of the fact that it contains fewer gene regulatory factors than other bacteria (Alm *et. al.*, 1999). Flagellar biosynthesis is a hierarchical process involving sequential activation of approximately 40 genes (Niehus *et al*., 2004; Macnab, 2003). However, in *H. pylori*, the master regulator of flagellar biosynthesis (FlhCD) has not been found which could link flagellar biosynthesis with the cell cycle (Spohn and Scarlatto, 1999; Niehus *et al*., 2004). In contrast to

other bacteria, the flagellar genes of *H. pylori* are scattered throughout the genome in 25 or more transcriptional units (Macnab, 1992). Flagellar biosynthesis studies in *H. pylori* have shown that regulation of flagellar genes is a complex process and involves three sigma factors (σ28, σ54, and σ80).

In *H. pylori*, the primary sigma factor (σ80) is required for the regulation of the expression of housekeeping genes (class I genes) along with some other essential genes. The other genes are the genes encoding flagellar export apparatus components, flagellar basal body proteins as well as chemotaxis proteins (Beier *et al*., 1997; Porwollik *et al*., 1999; Spohn and Scarlatto, 1999). σ54 (RpoN) and FlgR (σ54 dependent activator) are required for the expression of the genes (class II genes) encoding components of the hook, hook-associated proteins, the rod and the minor flagellin FlaB (Spohn and Scarlatto, 1999). In *Salmonella*, the hook length is under the control of FliK (Ferris and Minamino, 2006). In *H. pylori* FliK has been identified as a product of gene HP0906 (Ryan *et al*., 2005). The novel Zn-finger protein HP0958 interacts with the sigma factor RpoN (Rain *et al*., 2001). σ54 (RpoN) activity is modulated by completion of the hook structure in *H. pylori* (Douillard *et al*., 2008). A HP0958 deletion impairs transcription of the major flagellin (FlaA), minor flagellin (FlaB) and genes encoding flagellar components. σ28 (FliA) controls the expression of late flagellar genes (class III genes) including the gene encoding the major flagellin (flaA) as well as proteins involved in flagellin modification and export (flaG and fliS) (Niehus *et al*., 2004). It also regulates expression of the gene encoding the filament capping protein (fliD) (Kim *et al*., 1999; Leying *et al*., 1992; Tomb *et al*., 1997). σ28 activity is tightly regulated by the anti-$\sigma^{28}$ factor FlgM (Colland *et al*., 2001). FlgM secretion relieves this repression in other organism (Kutsukake, 1994) but in case of *H. pylori* FlgM secretion has not been demonstrated (Douillard *et al*., 2008).

### 1.3.3 Flagellum specific export proteins that are not evident in the *H. pylori* genome

Many flagellar export proteins and chaperones have been identified in *Salmonella*. Flagellar chaperones FlgN, FliT and FliS are required for the stability as well as of secretion of the substrates FlgK, FlgL, FliC and FliD (Auvray *et al*., 2001; Fraser *et al*., 1999; Bennett and Hughes, 2000). With the exception of a few flagellar genes, homologues for all of the export apparatus proteins have been found or annotated in *H. pylori* (Tsang and Hoover, 2014). The flagellar genes that have not been identified in *H. pylori* include export protein FliO, FliJ, FliT

and FlgN. This may be due to low sequence identity (Tsang and Hoover, 2014; Douillard *et al.*, 2010). Sequences of the proteins of flagellar export apparatus (FlhA, FlhB, FliP, FliQ, and FliR) and FliN have been found to be highly conserved among the various *Helicobacter* species. Out of the six flagellar export apparatus proteins, FliO is the least conserved protein. A potential FliO homolog, HP0583, have been characterized in *H. pylori* recently but it is predicted to have an N-terminal periplasmic domain which is absent in *Salmonella* FliO (Tsang and Hoover, 2014). Similar to FliO, a potential FliJ homolog (HP0256) have been identified by bioinformatics analysis using only the FliJ N-terminal coiled-coil domain as a search query (Douillard *et al* 2010). Although HP0256 has a different function in *H. pylori* than FliJ in *Salmonella*, it is definitely required for full motility. We will discuss in detail about HP0256 in section 1.4.3. FliT gene have been found in other bacteria like *E. coli, Salmonella, B. subtilis* (Chen and Helmann, 1994; Kawagishi *et al*., 1992) but in *H. pylori* fliT gene homologue has not been found (Tomb *et al*., 1997). Similarly, FlgN has not been reported in *H. pylori* (Lam *et al*., 2010).

## 1.4 Structures and functions of proteins FliH, FliI, FliJ, FliS and FlaA

### 1.4.1 FliH

FliH is one of the major soluble component of the flagellar export system and it interacts with the cytoplasmic component FliI (Minamino and Macnab, 2000b). Originally FliH was thought to be the regulator of FliI ATPase, since it reduces the ATPase activity of FliI potentially preventing futile ATP hydrolysis (Minamino and Macnab, 2000a). In other words, FliH acts as a regulator of FliI to prevent it from wasting ATP when FliI is not engaged in flagellar protein export (Minamino and Macnab, 2000b; González-Pedrajo *et al.*, 2002). However, overexpression of FliI ATPase or certain mutations in FlhA or FlhB can bypass the effects of FliH null mutants (Minamino *et al.*, 2003). Interestingly overexpression of FliH in otherwise wild-type cells also reduces motility which can be explained by the excessive formation of a FliI:$FliH_2$ heterotrimer (Minamino and Macnab, 2000b) thus apparently preventing FliI hexamerization and incorporation into the export apparatus. FliH binds to FliI and anchors the FliI hexamer to the export gate (Minamino *et al.*, 2014) thus FliH provides a link between FliI hexamer formation and protein export (Minamino *et al.*, 2003; Minamino and Macnab, 2000b; Lane *et al*., 2006). It is proposed that FliH plays an important role in the effective docking of

FliI to the export gate so that the energy of ATP hydrolysis can be usefully coupled to the translocation of export substrates (Minamino *et al.*, 2003).

In all, both FliI and FliH are required for efficient export, and they must be expressed in a proper ratio. The N-terminal and C-terminal domain sequences of FliH are homologous to those of the $F_OF_1$-ATP synthase, b and δ subunits, respectively (Pallen *et al.*, 2006; Lane *et al.*, 2006). *H. pylori* FliH consists of 258 amino acid residues while *Salmonella* FliH is 234 amino acid long. The *Salmonella* FliH sequence can be divided into three regions: an N-terminal region (residues 1–100: $FliH_N$), a central region (residues 101–140: $FliH_M$) and a C-terminal region (residues 141–235: $FliH_C$) (González-Pedrajo *et al.*, 2002) (Figure 1.8). *In vitro* characterization of *Salmonella* FliH has revealed that it forms an elongated dimer in solution (Minamino *et al.*, 2002; Lane *et al.*, 2006). A central region in *Salmonella* FliH (residues 101 and 140) is required for homodimer formation and contains a sequence with a probability of forming a helical coiled-coil. Residues 100–235 of *Salmonella* FliH are required for interaction with FliI. The N-terminal residues of FliH bind to the flagellar chaperone FliJ (Gonzàlez-Pedrajo *et al.*, 2002). A conserved AxxxG(xxxG)$_m$xxxA (where the value of m varies considerably) motif spanning residues 59–94 in *Salmonella* FliH suggests that these FliH GxxxG repeats may have a role in the regulation of the activity of FliI ATPase (Trost & Moore, 2009; Minamino *et al*., 2009). Deletion of residues 60-100 from *Salmonella* FliH have been shown to enhance FliI ATPase activity *in-vitro* (Gonzàlez-Pedrajo *et al.*, 2002). The extreme N-terminal region of FliH interacts with the transmembrane domain of FlhA (Noritaka *et al.*, 2012). Highly conserved residues of *Salmonella* FliH, Trp-7 and Trp-10, are not only important for interaction with FliN but also for docking of FliHI complex to the export gate thus it is suggested that the interactions of Trp-7 and Trp-10 with FliN allow the FliH–FliI complex to be localized to the C ring (Minamino *et al*., 2009). This interaction also increases local concentration of FliI near the export gate (Noritaka *et al.*, 2012). However, it is proposed that these interactions must be highly dynamic because localization of FliI or FliH to the C ring is one of the steps on its way to the docking platform (Minamino *et al.*, 2003; Minamino and Namba, 2008).

Recently the structure of *Salmonella* $FliH_C$ (residues 99-235) in complex with FliI has shown that $FliH_C$ forms unusually asymmetric homodimer (Imada *et al*., 2016). The two subunits ($FliH_C$-A and $FliH_C$-B) forming the $FliH_C$ dimer, have a different 3D arrangement in the FliHI complex (Figure 1.8 and Figure 1.9). $FliH_C$-A consists of a N-terminal helix (α1), a central

globular domain made up of two α-helices (α2 and α3) and four β-strands (β1-4) and a C-terminal helix (α4) (Figure 1.8). A small bend at residue Leu-119 divides α1 into two parts α1a and α1b. $FliH_C$-B consists of five α-helices (α1-4 and αs) and three β-strands (β1-3) (Figure 1.9). The N-terminal α1 helix is completely divided into two parts (α1a and α1b) by the kink between the two residues, Ile-123 and Ala-124. The hydrophobic core is exposed due to absence of α2 and β4 from the globular domain. The α2 helix turns its orientation while β4 changes its conformation to a loop. α4 adopts two completely different conformations in a dimer (Figure 1.9B).

$FliH_{C2}$ resembles the peripheral stalk subunits (E and G) of A/V type ATPases (Imada *et al*., 2016). Despite this information, the crystal structure of full length FliH (*Salmonella* / *H. pylori*) is not yet available. Crystal structures as well as careful biochemical/biological analysis of this protein should ultimately be able to solve the fascinating issues related to FliH structure and function.

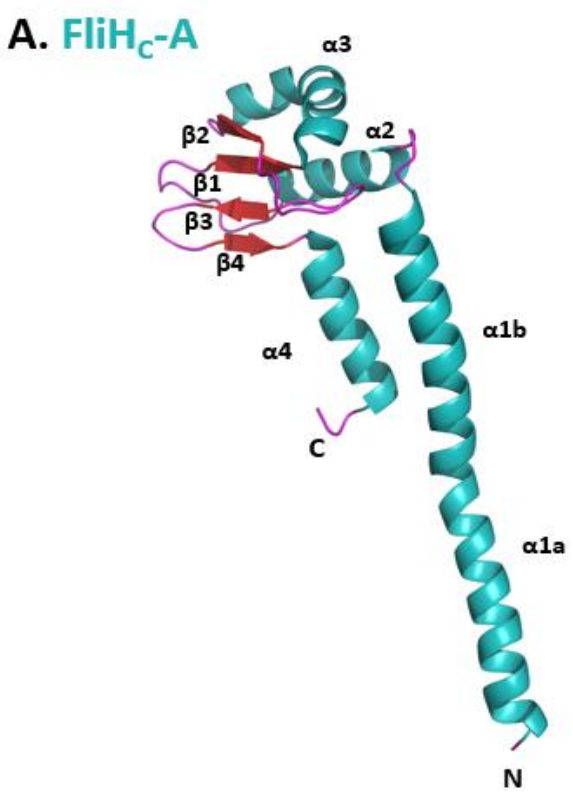


Figure 1.8. **Structure of *Salmonella* FliH$_C$-A (residues 99-235).**

The Cα ribbon drawing of FliH$_C$-A showing N-terminal helix (α1), a central globular domain (made up of two α-helices (α2 and α3) and four β-strands (β1-4)) and a C-terminal helix (α4). All the helices are shown in cyan and β-strands are in red. The figure was created using PyMOL using the structure of FliH$_{C2}$–FliI complex (PDB ID 5b0O) (Imada *et al.*, 2016).

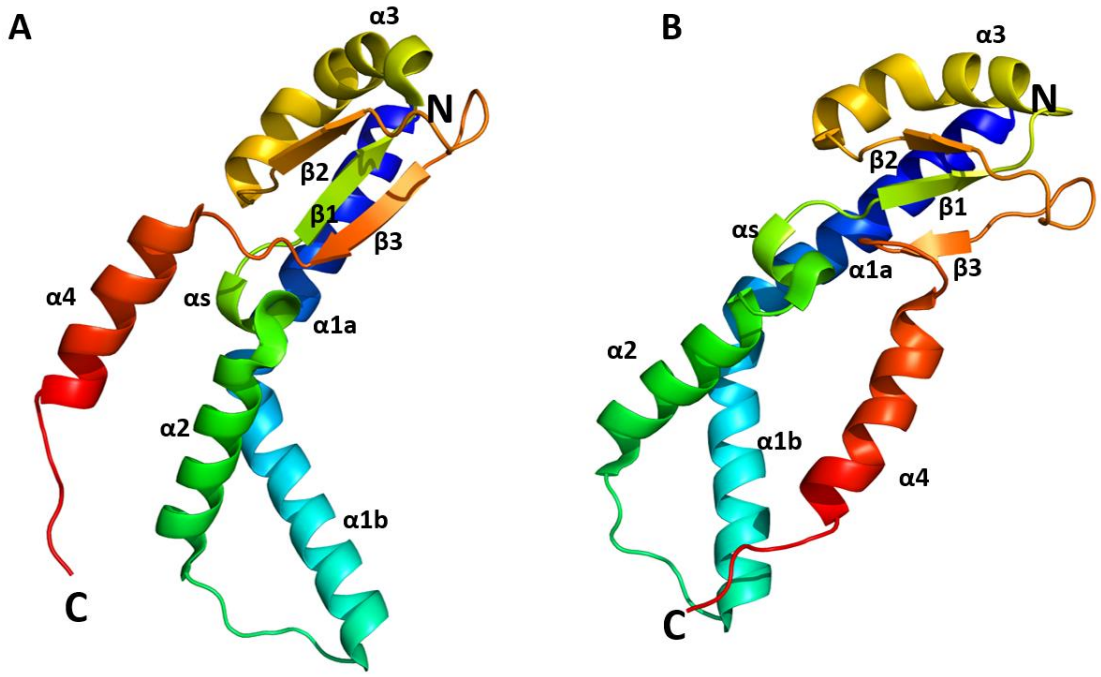


Figure 1.9. **Structure of *Salmonella* FliH$_C$-B (residues 99-235).**

The Cα ribbon drawing of FliH$_C$-B shown in rainbow colors (N-terminus (blue) to C-terminus (red)). FliH$_C$-B shows two different conformations in a dimer **A** and **B**. **A.** FliH$_C$-B consists of five α-helices (α1-4 and αs) and three β-strands (β1-3). The N-terminal α1 helix is completely divided into α1a and α1b. The hydrophobic core is exposed due to absence of α2 and β4 from the globular domain. The α2 helix turns its orientation while β4 changes its conformation to a loop. **B.** α4 completely changes its conformation. (Imada *et al.*, 2016). Both the figures were created using PyMOL using the structure of FliH$_{C2}$–FliI complex (PDB ID 5b0O).

### 1.4.2 FliI

The soluble component FliI of the flagellar export system is a Walker type ATPase (Fan *et al.,* 1996). FliI was thought to convert chemical energy from ATP hydrolysis into mechanical work which drives the export process since FliI containing mutants show an impaired mechanism of flagellar protein export (Fan *et al.,* 1996 and Jenks *et al.,* 1997). FliI shows extensive similarity to the α and β subunits of $F_OF_1$ ATP synthase (Imada *et al.,* 2007) but this sequence similarity was mostly limited to the ATPase domains (Lane *et al.,* 2006 and Volger *et al.,* 1991). FliI self assembles to form a homohexamer (Auvray *et al.,* 2002; Claret *et al.,* 2003), especially in the presence of ATP or FliJ. The ATPase InvC from T3SS shares a high degree of sequence identity with FliI. InvC has a similar role of substrate recognition and the hexameric form of InvC is required for higher efficiency (Roblin *et al*., 2014).

*H. pylori* FliI is a 47 kDa (434 amino acids) protein. FliI consists of three domains - the N-terminal domain, which interacts with FliH (Minamino *et al.*, 2000, Lane *et al*., 2006), the central domain showing the presence of a catalytic site for ATP hydrolysis, and a helical C-terminal domain (Imada *et al.,* 2007). The catalytic site is conserved along with the residues of the P-loop (Walker A motif), where phosphates of the ATP binds. *H. pylori* FliI shows 26% sequence identity with the β-subunit of the bovine $F_1$ - ATPase. The structure of *Salmonella* FliI shares a striking structural similarity to α and β subunits of $F_1$ - ATPase (Imada *et al.,* 2007). Residues of α & β subunits involved in ATP hydrolysis are highly conserved in FliI (Imada *et al.,* 2007 & Vogler *et al.,* 1991).

An N-terminal double mutant (R7C/L12P) or the mutant lacking first seven residues of *Salmonella* FliI, failed to interact with FliH *in vitro* (Minamino *et al.,* 2000). N-terminal truncated versions of FliI do not form a complex with FliH hence the FliI N-terminus is thought to play an important role in FliH interaction (Minamino *et al.,* 2001; Lane *et al.,* 2006). The sequence of the central ATPase domain shows extensive similarity to the primary sequence of the β subunit of $F_1$-ATPase (Vogler *et al.,* 1991) (Figure 1.10). It was thought that FliI and $F_1$-ATPase share a similar pathway for ATP hydrolysis since ADP binds to the P-loop (Walker-loop) of FliI in a similar fashion to that in the $F_1$ ATPase α and β subunits (Imada *et al.,* 2007).

Previously it was suggested that *Salmonella* FliI self-assembles into a homo-hexameric ring and exerts its ATPase activity (Claret *et al.,* 2003; Minamino *et al.,* 2006). In solutions

lacking ATP, FliI is largely monomeric while the amount of the FliI hexamer increases with increase in the ATP concentration or the concentration of its non-hydrolysable analog, ATPγS or AMP-PNP (Adenosine 5'-(β, γ-imido) triphosphate tetra lithium salt hydrate) (Claret *et al.*, 2003). From *in vitro* reconstitution experiments, it was previously observed that the binding of $Mg^{2+}$-ATP or analogs of ATP to FliI induce $FliI_6$ ring formation and that ATP hydrolysis followed by release of ADP and Pi destabilizes the ring (Minamino *et al.*, 2009). Hence it was suggested that the assembly-disassembly cycle of the FliI hexameric ring occurs is an ATP dependent process. In agreement with this suggestion, the K188I mutation of *Salmonella* FliI, which reduces the affinity of FliI for ATP, remarkably decreased the efficiency of FliI hexameric ring formation (Bai *et al.*, 2014). On the other hand, the E211Q active site mutation of *Salmonella* FliI, resulting in the complete loss of the ATPase activity, formed FliI rings in the presence of $Mg^{2+}$-ATP (Kazetani *et al.*, 2009). This suggests that the assemble-disassembly of FliI hexamer occurs in ATP-dependent manner.

FliI forms a complex with a homo-dimer of FliH (Minamino *et al.*, 2000). It has been proposed that the $FliH_2FliI$ complex acts as a dynamic carrier and the $FlhA_9$ ring acts as a sorting platform. FlhA allows fast and efficient entry of export substrates into the export gate with the help of the FliI hexameric ring, which appears to act as a static substrate loader (Bai *et al.*, 2014). Recently, it has been indicated that the rate of ATP hydrolysis by FliI does not directly determine the flagellar protein export and thus the FliI hexameric ring does not facilitate the entry of substrates into the export gate (Minamino *et al.*, 2014).

Figure 1.11 shows the 2.4 Å resolution structure of *Salmonella* FliI (Δ1–18) in the presence of ADP (Imada *et al.*, 2007). Removal of the first 18 residues from native FliI made it more soluble and easy for crystallization. The asymmetric unit of the crystal contains two molecules of FliI (Δ1–18) and these molecules were related by a pseudo 2-fold symmetry axis. The N-terminal domain contains six β strands while the ATPase domain having 272 residues shows an α / β fold. The ATPase domain is composed of 10 α helices and13 β strands. The central part of the ATPase domain has a P loop containing the Walker A and B motif, (Figure 1.11). The C-terminal domain has two loops that connect three α helices. The ATPase domain is connected to N-terminal domain through a long linker segment (Asn-98 through Lys-108).

**FliI active site mutation**

Glutamate 193 plays an important role in ATP hydrolysis and it is conserved in FliI homologs. The carboxyl group of Glutamate 193 (similar to Glu-190 in the thermophilic *Bacillus* $F_1$ - ATPase), polarizes a water molecule for the nucleophilic attack to the γ-phosphate of ATP (a general base catalysis), and hence the replacement of Glutamate by Glutamine should result in the complete loss of ATPase activity (Shimabukuro *et al.*, 2003; Kazetani *et al.*, 2009) since the carboxyl group will be now replaced by an amide group.

The E211Q mutation in *Salmonella* FliI caused a loss-of-function phenotype and it significantly reduced the ATPase activity of FliI (Kazetani *et al.*, 2009). FliI activity increases in the presence of ATP/FliJ, probably due to allosteric activation of the enzyme. Thus we hoped that the expression of *H. pylori* FliI E/Q mutant in the presence of AMP-PNP (a non-hydrolysable analog of ATP) will help to reveal whether this substrate binds to the active site of the FliI which might create a conformational change thus regaining ATPase activity allosterically. It is shown that the yield of complete hexameric FliI rings increases by 20-fold when AMP-PNP was used instead of ATP, presumably by stabilizing the assembled oligomers (Claret *et al.*, 2003), hence we hope to determine the structure of FliI ATPase in the presence of AMP-PNP.

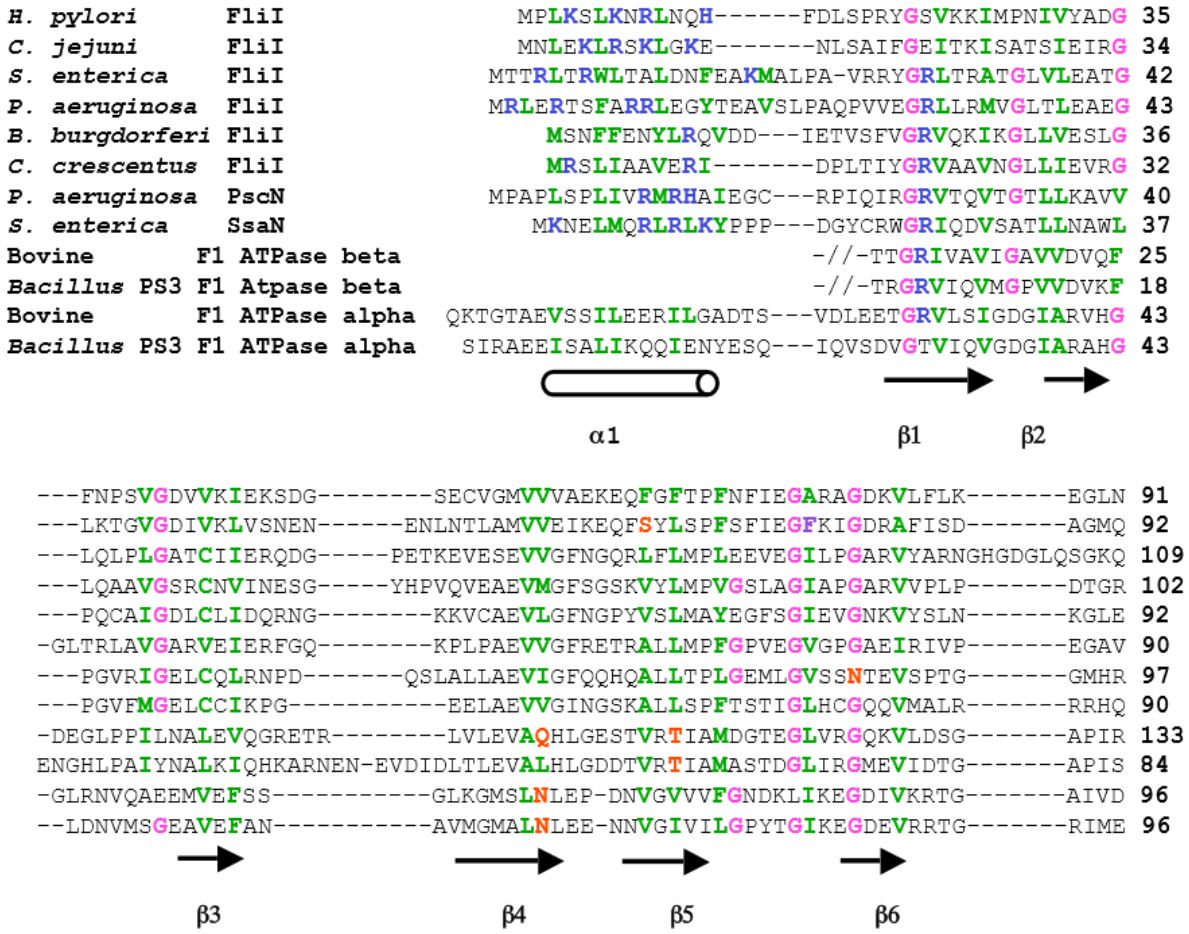


Figure 1.10. **A multiple sequence alignment of the N terminus of FliI, Type III secretion ATPases from various bacteria, and α and β subunits of $F_1$-ATPase.**

Secondary structure units are shown for the known structures of the N-terminal domains of $F_1$ ATPase α and β subunit. The α-helices are depicted by cylinders and β-strands are depicted by arrows. Residues making up the hydrophobic core in the $F_1$ ATPase subunits are shown in green. Highly conserved residues are shown in pink. *H. pylori* FliI shows 48% sequence identity with *Salmonella* FliI. (Figure from Lane *et al*., 2006).

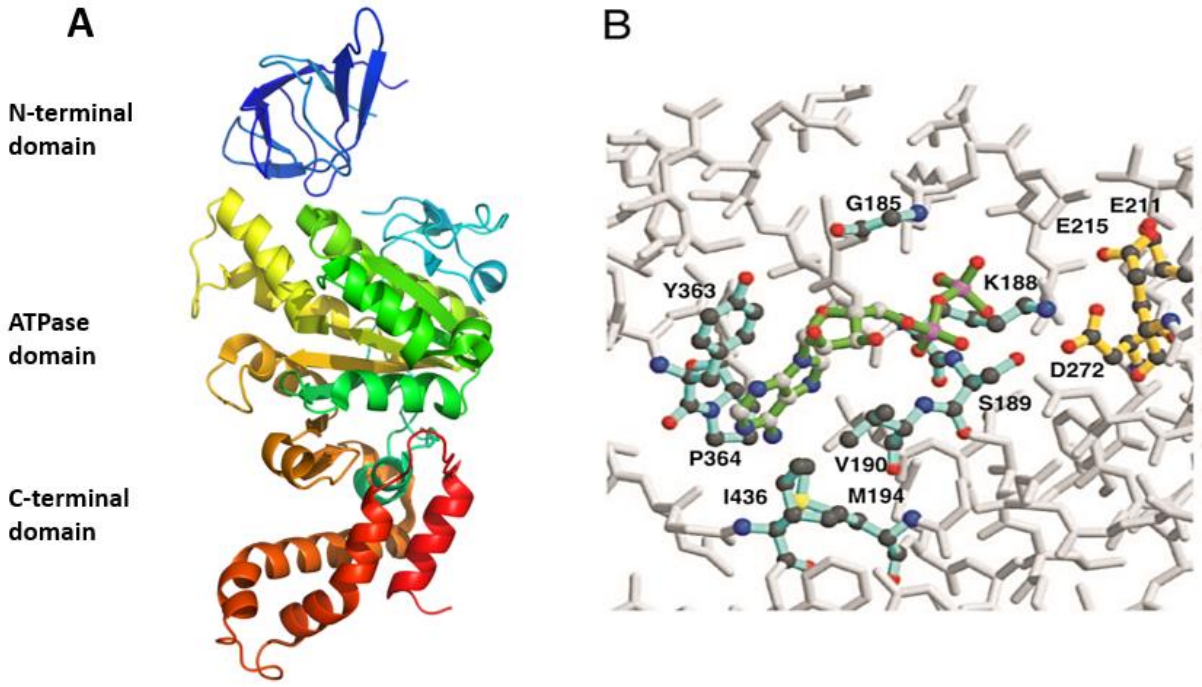


Figure 1.11. **Structure of *Salmonella* FliI (Δ1–18).**

A. The structure of FliI (Δ1–18) showing N-term β-sheet in blue color, central α-helical ATPase domain in green and C-terminal domain in red. The missing linker region (from the model) that connects the N-terminal and ATPase domain is shown by dashed line. B. The stereo view of the nucleotide-binding site showing ADP is green, the residues interacting with ADP are colored cyan. The conserved residues taking part in catalysis are shown in yellow. (Imada *et al.*, 2007). Figure was created using PyMOL.

### 1.4.3 FliJ

Out of the three soluble components of the flagellar export system, the third protein is FliJ. It is a small (147 residue) protein in *Salmonella*, with a deduced molecular mass of 17 kDa (Ibuki *et al.,* 2008). *Salmonella* FliJ has a high proportion (31 mol %) of charged residues, and a high-helical propensity (Vogler *et al.,* 1991). Many Type III export chaperones also show these characteristics (Bennet and Hughes, 2000). FliJ is an essential component for protein export in *Salmonella,* as lack of FliJ results in a leaky motile phenotype (Minamino *et al.,* 2000). It interacts with various flagellar axial proteins. Originally FliJ was postulated to function as a general chaperone for the export substrate proteins (Minamino *et al.,* 2000) but a recent structural study established that it is homologous to the γ subunit of the $F_OF_1$ ATP synthase (Ibuki *et al.,* 2011) (Figure 1.12) and hence may function as an integral component of the FliH–FliI–FliJ ATPase complex. FliJ binds to other export chaperones like FliT and FlgN (Evans *et al*., 2006), hence may function as a co-chaperone.

FliJ interacts not only with the FliH–FliI complex but also with the cytoplasmic domains of FlhA and FlhB (export-gate proteins) (Minamino & Macnab, 2000a; Fraser *et al.,* 2003). FliJ modifies the binding abilities of $FlhA_C$ (Bange *et al.,* 2010). It also interacts with FliM and other flagellar specific chaperones such as FliT and FlgN (Gonza´lez-Pedrajo *et al.,* 2006, Evans *et al.,* 2006). FliJ was shown to have a moonlighting role to cycle export chaperones FlgN (chaperone for hook–filament junction proteins FlgK and FlgL) and FliT (chaperone for filament cap protein FliD), but not FliS, the chaperone of the major filament protein, FliC (flagellin) (Evans *et al.,* 2006). Overexpression of FliJ in otherwise wild-type cells also reduces motility (Minamino *et al.,* 2000a). FliJ stabilizes FliI hexamric ring (Ibuki *et al*., 2011).

FliJ is considered to regulate the export process through complex interactions, but the detailed function of FliJ is still unknown. Therefore, the three-dimensional structure of FliJ is essential in order to understand the mechanism of flagellar protein export as well as the role of FliJ in export process. In the case of *H. pylori* there is no strong evidence for the presence of FliJ. Previous bioinformatics, motility and mutagenesis studies have identified HP0256 as a potential FliJ homologue (Douillard *et al*., 2010). HP0256 from *H. pylori* shows 38% sequence identity with *Salmonella* FliJ (Figure 1.13). HP0256 gene encodes a 17 kDa protein having pI of 7.8. *B. subtilis* FliJ is predicted to have α-helical structure with the possibility of forming a C-terminal coiled-coil (Wolf *et al.,* 1997).

The structure of *Salmonella* FliJ (PDB ID 3AJW) (Ibuki *et al*., 2011) (Figure 1.12B) showed that FliJ forms an antiparallel coiled-coil structure made up of two helices, α1 and α2. Since α1 was shorter than α2, the N- and C-terminal short segments of α2 protruded from the core of coiled-coil. Due to strong aggregating tendacy of FliJ, it was difficult to crystllize hence a His-tag was attached to N-terminus which made it more soluble and easy to crystallize (Ibuki *et al*., 2011).

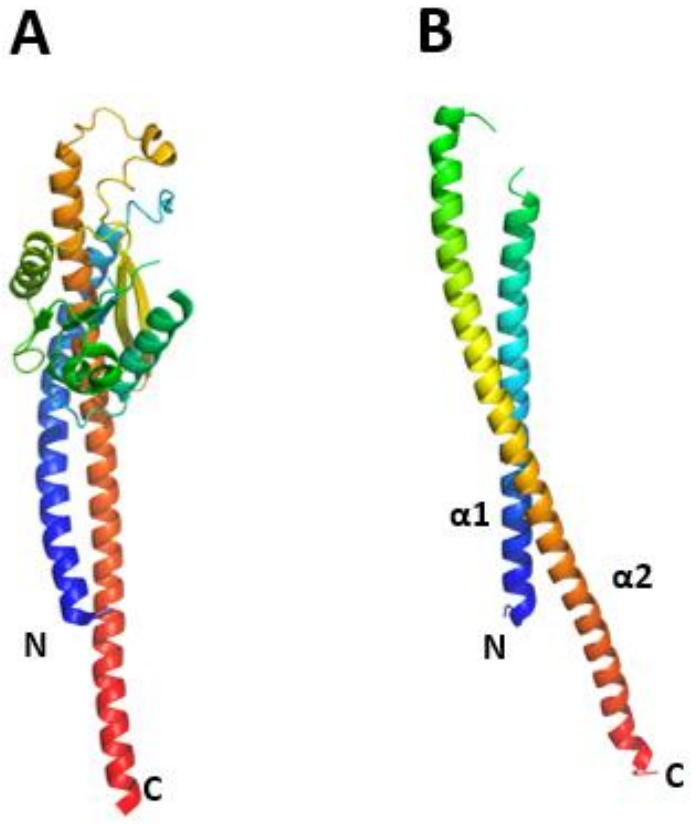


Figure 1.12. **Structural similarity between *Salmonella* FliJ and γ subunit of $F_1$-ATPAse.**

**A**. Cα ribbon diagram of the γ subunit of bovine $F_1$-ATPase (PDB ID 1E79). **B**. Cα ribbon diagram of *Salmonella* FliJ (PDB ID 3AJW). The binding regions for FliH (red-orange), FliT and FliH both (yellow), FliT (green) FlgN (sky-blue) respectively (Ibuki *et al*., 2011). Both the figures were created using PyMOL (Ibuki *et al*., 2011).

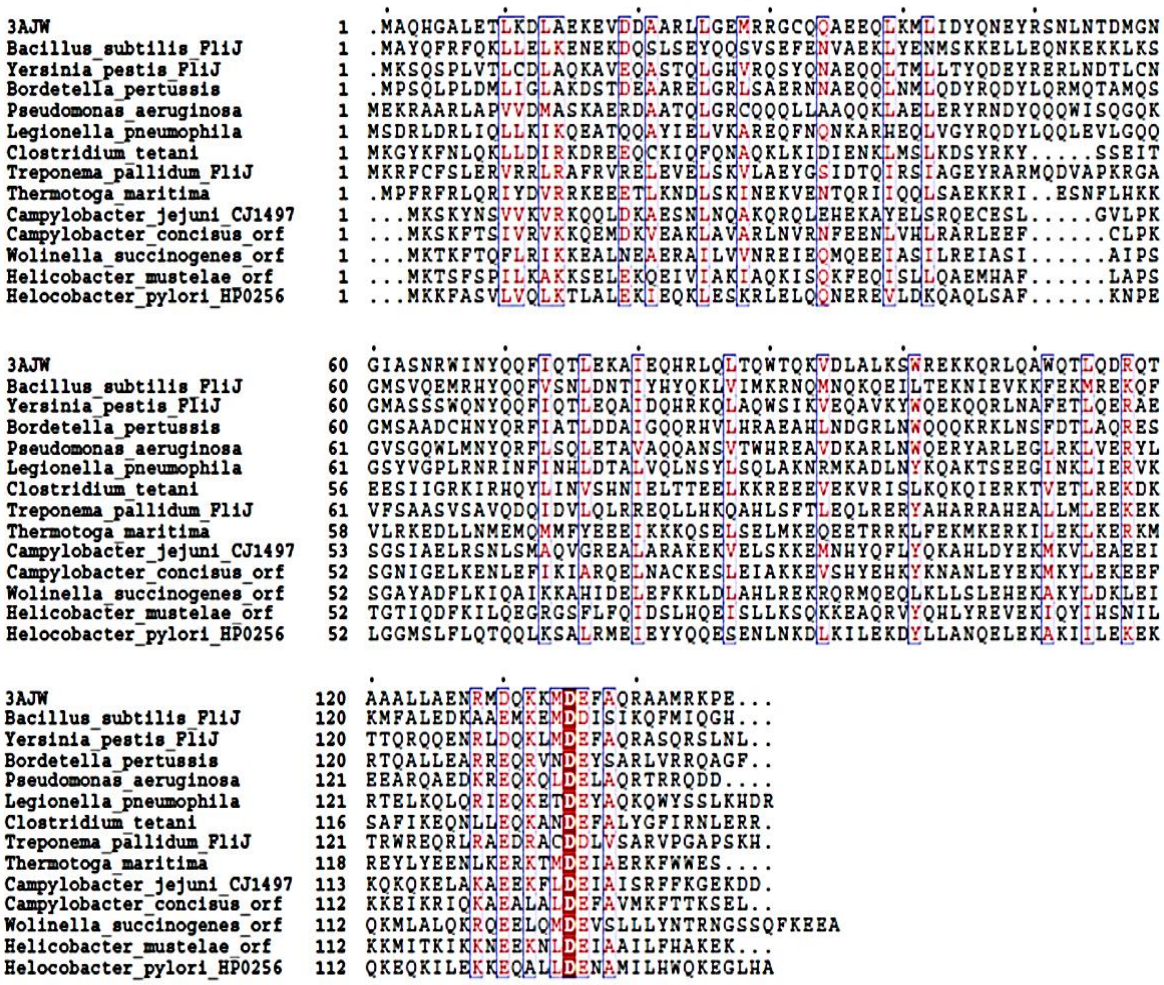


Figure 1.13. **Sequence similarity between FliJ homologs.**

A multiple sequence alignment of the HP0256 and FliJ homologs from other organisms. The conserved residues are indicated in red. HP0256 shows highest sequence identity with *Salmonella* FliJ (3AJW) and *C. Jejuni* FliJ whereas lowest sequence identity with *C. tetani*. HP0256 from *H. pylori* shows 38% sequence identity with *Salmonella* FliJ.

### 1.4.4 The FliH:FliI:FliJ complex

FliI ATPase forms a hexamer to fully exert its ATPase activity (Claret *et al*., 2003) and FliJ forms the $FliI_6FliJ$ complex by binding to the center of FliI hexameric ring (Ibuki *et al*., 2011). FliI also forms a complex with a homo-dimer of FliH (Imada *et al*., 2007). The $FliH_2FliI$ complex is believed to deliver export substrates as well as chaperone-substrate complexes to the export gate because chaperone-substrate complexes bind to the $FliH_2FliI$ complex through cooperative interactions among FliI, chaperone, and substrate (Thomas *et al*., 2004; Minamino *et al*., 2012). An interaction of the extreme N-terminal region of FliH with FliN allows the $FliH_2FliI$ complex to bind to the C ring (Minamino *et al*., 2009). FliJ interacts with FliI hexameric ring and with the $FliH_2FliI$ complex (González-Pedrajo *et al.*, 2002). It is proposed that FliH helps the FliJ and the chaperone-export substrate complexes to anchor to its docking platform ($FlhA_9$-$FlhB_C$) (Bai *et al*., 2014). Taken together, FliI and FliH, both are required for delivery of export substrates to the export gate (Minamino *et al*., 2008a) and it is proposed that FliH and FliI helps to enforce the interaction between FlhA and its binding proteins to maintain an efficient energy coupling mechanism of flagellar protein export (Minamino *et al*., 2016)

Recently, the structure of *Salmonella* $FliH_C$-FliI heterotrimer has been solved (Imada *et al.*, 2016) which showed that $FliH_{C2}$ forms an unusually asymmetric homodimer that binds to the N-terminal domain of FliI (Figure 1.14). The Structure of $FliH_C$-A and $FliH_C$-B is shown in results section 1.4.1. The globular domain of $FliH_C$-A interacts with the N-terminal α-helix of FliI. This helix fits into the hydrophobic groove made up of α1 helices of both $FliH_C$. The residues from *Salmonella* FliI that are important for this interaction are Arg-4, Arg-7, Trp-8, Leu-12, Phe-15 and Met-19. The hydrophobic grove is formed by residues Ile-123, Leu-127, Met-128, and Leu-132 (from α1b domain of $FliH_C$-A) and Phe-112, Leu-116, and Leu-119 (from α1a domain of $FliH_C$-B), Leu-127, Met-130, and Ala-134 (from α1a domain of $FliH_C$-B).

In this crystal structure, the interaction of $FliH_C$ with the oligomerization interface of FliI is not observed hence it is still not clear how FliH regulates the FliI hexamerization. In A/V type ATPases the conformational change of peripheral stalk subunit is required for assembly of the complex (Bernal and Stock, 2004) thus it might be possible that the change in conformation of FliH affects the FliI ring formation.

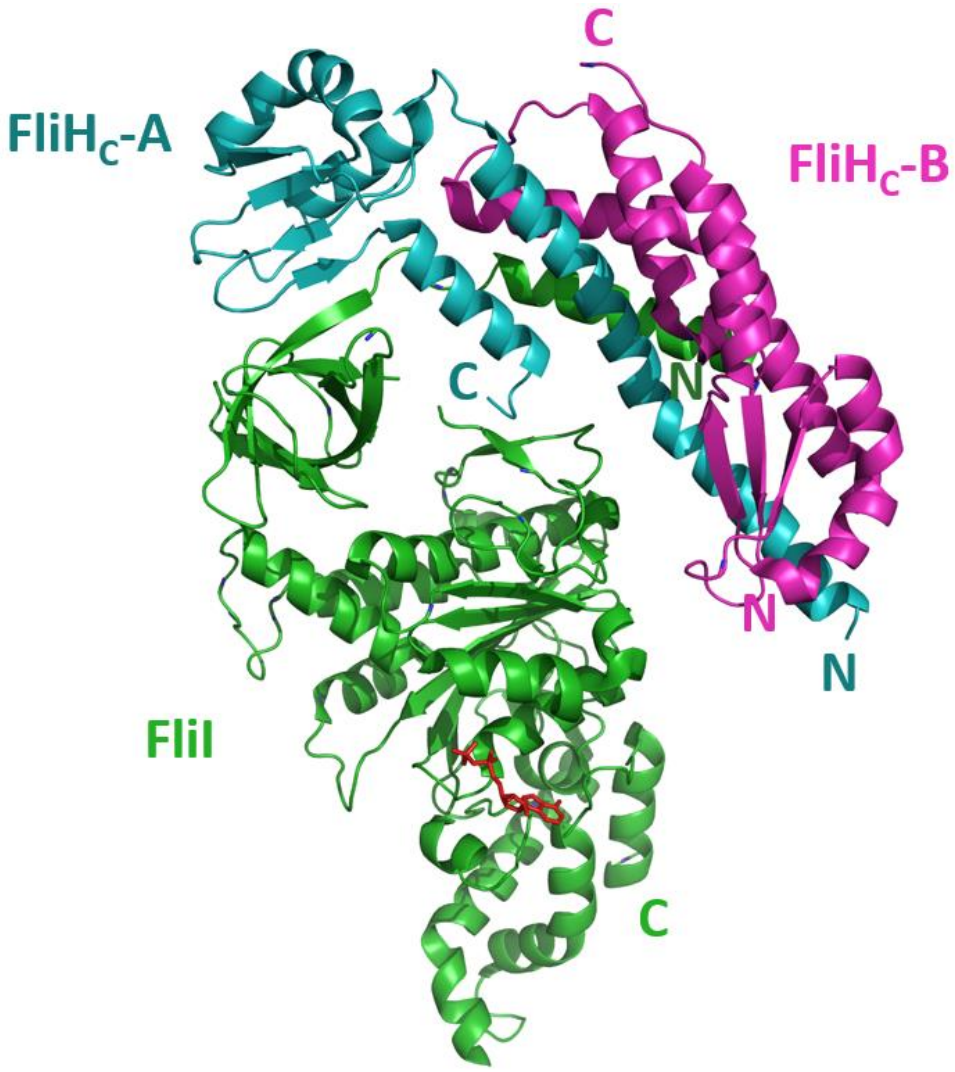


Figure 1.14. **The structure of *Salmonella* FliH$_{C2}$–FliI.**

Cα ribbon representation of the FliH$_{C2}$–FliI complex showing FliI (green) and two FliH$_C$ subunits (FliH$_C$-A and FliH$_C$-B in cyan, and magenta, respectively). The ADP molecule is shown in red. The figure was created using PyMOL using the structure of FliH$_{C2}$–FliI complex from (Imada *et al.*, 2016) (PDB ID 5b0O).

### 1.4.5 FliS

Studies of the flagellar export system have shown that FliS acts as the flagellin-specific chaperone that binds to flagellin to prevent premature polymerization of newly synthesized flagellin molecules (Auvray *et al*., 2001; Ozin *et al*., 2003; Muskotàl *et al*., 2006). From the genetic analyses in various bacteria like *E. coli* (Kawagishi *et al*., 1992), *Bacillus* (Chen and Helmann, 1994), and *Salmonella* (Yokoseki *et al*., 1995), it was confirmed that the *fliS* gene is necessary for flagellation. FliS also acts as a negative regulator for the export of the flagellum-specific anti- sigma factor, FlgM, in *Salmonella* (Yokoseki *et al*., 1996) but it is not yet clear whether it plays the same role in other microbes. In *H. pylori*, mutation in FliS results in non-motile and non-flagellate bacteria (Allan *et al*., 2000) due to reduced levels of both flagellin as well as the flagellar hook subunit FlgE suggesting that FliS might be involved in export of hook component proteins.

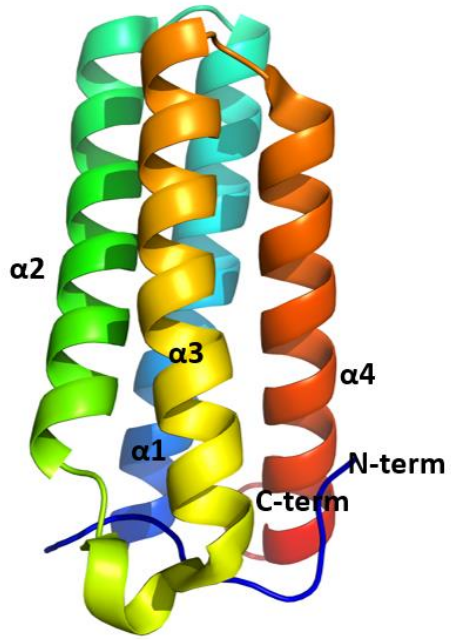


Figure 1.15. **Cartoon representation of structure of FliS from *H. pylori*.**

The structure of *H. pylori* FliS (PDB ID 3IQC) showing the presence of a core 4-helix bundle. The figure was drwan using PyMOL (Lam *et al*., 2010).

In addition to the flagellin subunits (FlaA and FlaB), FliS has shown to interact with a number of other proteins (Lam et al., 2010). *H. pylori* FliS consists of 126 amino acid residues and it shares 37% sequence identity with *Salmonella* FliS. The crystal structure of FliS from *Aquifex aeolicus* showed that it has a 4-helix bundle structure. This 4-helix bundle interacts with the C-terminal disordered region of flagellin which is neccessary for oligomerization. The structure of *H. pylori* FliS solved at 2.7 Å showed that FliS consists of a 4-helix bundle (Lam et al., 2010) (PDB ID 3IQC) (Figure 1.15). The flagellin binding pocket is present near the N-terminal end and this pocket is blocked by conserved residues Tyr10, Tyr28 and, Tyr 84. Residues of the surface loop between helices showed low sequence identity with FliS from *A. aeolicus*. Tyr10 occupied position in the hydrophobic pocket similar to FliS from *A. aeolicus* suggesting that FliS-flagellin must be conserved in other organisms like *H. pylori*. It also suggests that N-terminal fragment is important for blocking of flagellin binding. The exact mechanism of action of FliS in *H. pylori* remains to be elucidated.

### 1.4.6 Flagellin

A major flagellin subunit, FlaA (53 kDa) and a minor flagellin subunit FlaB (54 kDa) are the two structural subunits of *H. pylori* flagellar filament (Josenhans and Suerbaum, 2002; Leying *et al.*, 1992). The genes encoding both these flagellins are located at different sites on the *H. pylori* chromosome and are transcriptionally regulated by distinct promoters (Josenhans and Suerbaum, 2002). σ factors play an important role in controlling transcription of multiple flagellin genes (Chevance and Hughes, 2008). σ 70, σ 28, and σ 54 are mostly involved in recognition of flagellar gene promoters. In most of the bacteria containing multiple-flagellins, the major flagellins are under the control of σ 28, such as in *Helicobacter pylori*, *Salmonella etc*. (reviewed in Sun *et al*., 2014). Both of these flagellins play essential role in persistent colonization of the bacterium (Eaton *et al*., 1996). However, one of the flagellins (FlaA) constitutes the major part of the flagellar filament.

### 1.4.7 FliS:Flagellin complex

The *B. subtilis* FliS-flagellin complex binds to the cytoplasmic domain of FlhA (one of the export gate component) (Bange *et al*., 2010). The structure of *H. pylori* FlhA has been already solved in Dr. Moore's lab (Moore and Jia, 2010). We hypothesized that in *H. pylori*, the FliS:Flagellin complex should interact with FlhA. Therefore, in order to study the interaction of

FliS-flagellin with FlhA, we first decided to understand FliS-flagellin interaction in *H. pylori*. Structural studies in *A. aeolicus* have demonstrated that FliS forms a complex with flagellin in 1:1 stoichiometry (PDB ID 1ORY). In *A. aeolicus*, the C-terminal residues 479-518 of flagellin (FliC) were found to wrap around the surface of FliS in a horseshoe-like confirmation (Figure 1.16A). The hydrophobic interactions are predominantly present on the interface between the two molecules. The binding of FliS by flagellin is accompanied by the conformational changes in the FliS molecule. The N-terminus of FliS is disordered and it forms a quasi-helical cap. This cap is loosely bound and upon binding of flagellin it gets rearranged to form a short helix on the side of FliS helix bundle. This new conformation of FliS has a groove for binding of one of the α-helical segments (residues 510–518) of FliC. At the same time, the helical cap of FliS is replaced by the adjacent helical (residues 499–505) segment of FliC. The residues Ala504–Gln505 of FliC occupy equivalent positions as the residues Ala7–Tyr8 of FliS before binding (Figure 1.16).

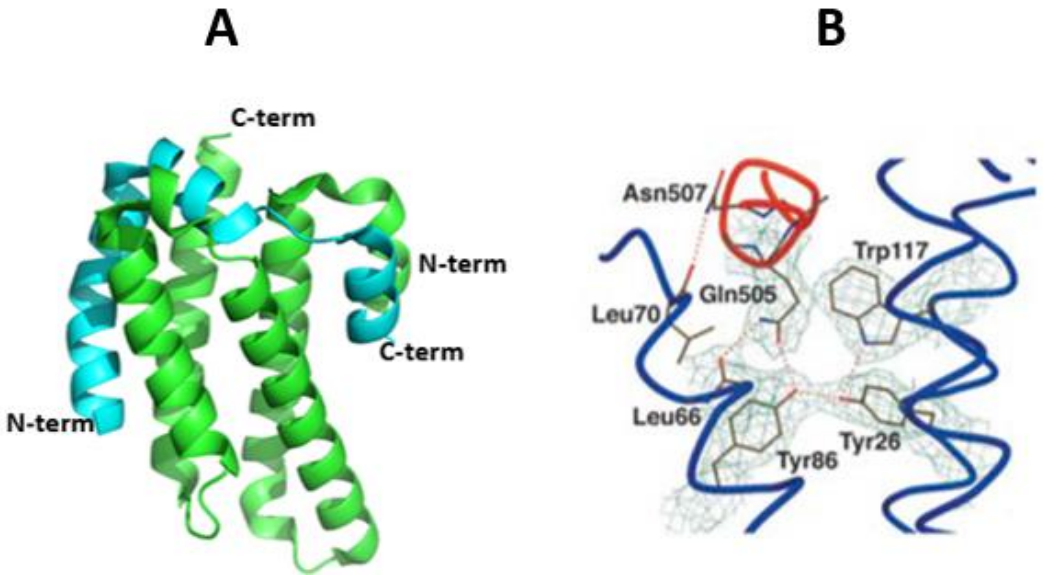


Figure 1.16. **The structure of *A. aeolicus* FliS-FliC complex.**

A. The structure of *A. aeolicus* FliS (green) (1002-1121) in complex with FliC (464-518) (orange). The C-terminal end residues of FliC interact with FliS and are wrapped around the structure of FliS showing horseshoe like confirmation. Figure was created using PyMOL (PDB ID 1ORY). B. The intramolecular interactions in the hydrophobic pocket of FliS showing Tyr8. Tyr8 of FliS and Gln505 of FliC occupy same position in the hydrophobic pocket of FliS. Hydrogen bonds are shown as dashed lines. Figure adapted from Evdokimov *et al*., 2003.

Therefore, the FliS appears to use the ‘stopper’ mechanism wherein the N terminus of FliS acts as a ‘molecular stopper’ to close the hydrophobic binding site whenever FliS is not bound to FliC. The high degree of conservation among the residues of FliS that form the binding pocket for Gln505 of FliC suggests that Gln505 is a key factor for binding. The high density of intermolecular contacts in this region also suggest that Gln505 is important determinant of the binding.

## 1.5 Project hypothesis and objectives

### Hypothesis

I. The structure of a FliH C-domain fragment will explain how FliH dimerizes and will reveal the tertiary structure of the FliH C-domain globular fragment (residues 171-258). Finally, the structure of the FliH C-domain will provide the framework for understanding how FliH binds to FliI. A FliI active site mutation should inactive the ATPase activity which should reveal the link between FliI hexamer assembly and disassembly cycle. This will help understand the mechanism of export.

II. Structural and biochemical information of HP0256 will help to identify HP0256 as a homolog of FliJ in *H. pylori*.

III. The *H. pylori* FliS interacts with flagellin. This interaction study will help to understand about how FliS:FliC interact with FlhA.

### Objectives

I. Cloning, purification and crystallization of FliH and FliI for determination of the structural basis for FliH dimerization as well as its interaction with FliI ATPase.

II. Cloning, purification and crystallization of HP0256 to confirm the presence of FliJ in *H. pylori* by determining the structural details of HP0256.

III. Protein characterization and crystallization studies of FliS to elucidate the structure of FliS and understand the basis of its interaction with the flagellin.

## 2. Materials and Methods

### 2.1 Primers used for cloning of *H. pylori* FliH (full length), FliH (57-258), FliH (73-258), FliI E193Q mutant, HP0256, FliS, $FlaA_1$ (411-510) & $FlaA_2$ (465-510)

Table 2.1. **List of the PCR primers used in this study for cloning of full length *H. pylori* NCTC 11637 FliH (full length), FliH (57-258), FliH (73-258), FliI E193Q mutant, HP0256, FliS, FlaA1 (411-510) & FlaA2 (465-510).**

Each 5' primer contains a *Bam*HI site and each 3' primer contains an *Eco*RI site for directional cloning into the pGEX-6P-3 GST fusion protein expression vector. The restriction sites as well as the mutagenic triplet codons with desired single amino acid substitution is shown in boldface italic.

| **Name of primer** | **Sequence** |
|---|---|
| **ML076** (Forward primer for FliH 2-258) | 5'- GACCTG***GGATCC***TCATTGAATTAGCCGTAAGAACTT GATTC – 3' |
| **ML021** (Reverse primer for all genes) | 5'-GGTTT***GAATTC***TCACACCTTAAAATTTTCCAACAC – 3' |
| **ML095** (Forward primer for FliH 73-258) | 5'- GACCTG***GGATCC***AGCCATTTAGTGAAATTGCAAATGCAG – 3' |
| **ML096** (Forward primer for FliH 57-258) | 5'- GACCTG***GGATCC***GAAACGATTTGATTGATTGCTTTGTTG – 3' |
| **ML097** (Forward primer for FliIE193Q mutant) | 5'-GTGATCGCTTTGATTGGG***CAG***AGGGGCAGAGAAATCCCT-3' |
| **ML098** (Reverse primer for FliIE193Q mutant) | 5'-AGGGATTTCTCTGCCCCT***CTG***CCCAATCAAAGCGATCAC-3' |
| **ML099** (Forward primer for $FlaA_1$ (411-510)) | 5'-GACCTG***GGATCC***GGATCTGGGGTT ACAACCTTG-3' |
| **ML100** (Forward primer for $FlaA_2$ (465-510)) | 5'-GACCTG***GGATC***CGCTGAGTCTAA ATCAGGGATGTG G-3' |

| | |
|---|---|
| **ML101** (Reverse primer for $FlaA_1$ and $FlaA_2$**)** | 5'-GACGAG***GAATTC***CTAAGTTAAAAGCCTTAAGATATTTTGTTG-3' |
| **ML102** (Forward primer for FliS) | 5'-GACCTG***GGATCC***ATGCAATACGCTAACGCTTATC-3' |
| **ML103** (Reverse primer for FliS) | 5'-GACGAG***GAATTC***TTAGGCGAGTTCATCTGAATGGATTTC-3' |
| **ML104** (Forward primer for HP0256) | 5'-GACCTG***GGATCC***AAAAAATTCGCTTCTGTGTTGGTG-3' |
| **ML105** (Reverse primer for HP0256) | 5'-GACGAG***GAATTC***TTACGCATGCAAGCCCTC TTTTTGCC-3' |

## 2.2 Cloning procedures

### 2.2.1 DNA amplification by PCR

The standard cloning protocol for the cloning of gene fragments by Polymerase Chain Reaction (PCR) is as follows. PCR was carried out in a MyCycler$^{TM}$ Thermal Cycler (Bio-Rad) in a total reaction volume of 100 µL. Each reaction contained 1 ng/µL of template DNA, 1X reaction mixture (Fermentas) (10X *Pfu* buffer + MgSO4), 0.5 mM of dNTP, 200 nM of forward and reverse primers, 1 µL of (5000 unit) *Pfu* DNA polymerase (Fermentas). A typical amplification program (Figure 2.1) consists of total 30-35 cycles. Each cycle consisted of 1 min incubation at 95°C for denaturation of DNA, a 45 sec annealing step at 55°C and 2 min elongation

step at 72°C. In the first cycle, the denaturation step was extended up to 2 min. In the last cycle, the elongation step was extended to 10 min before final hold at 4°C.

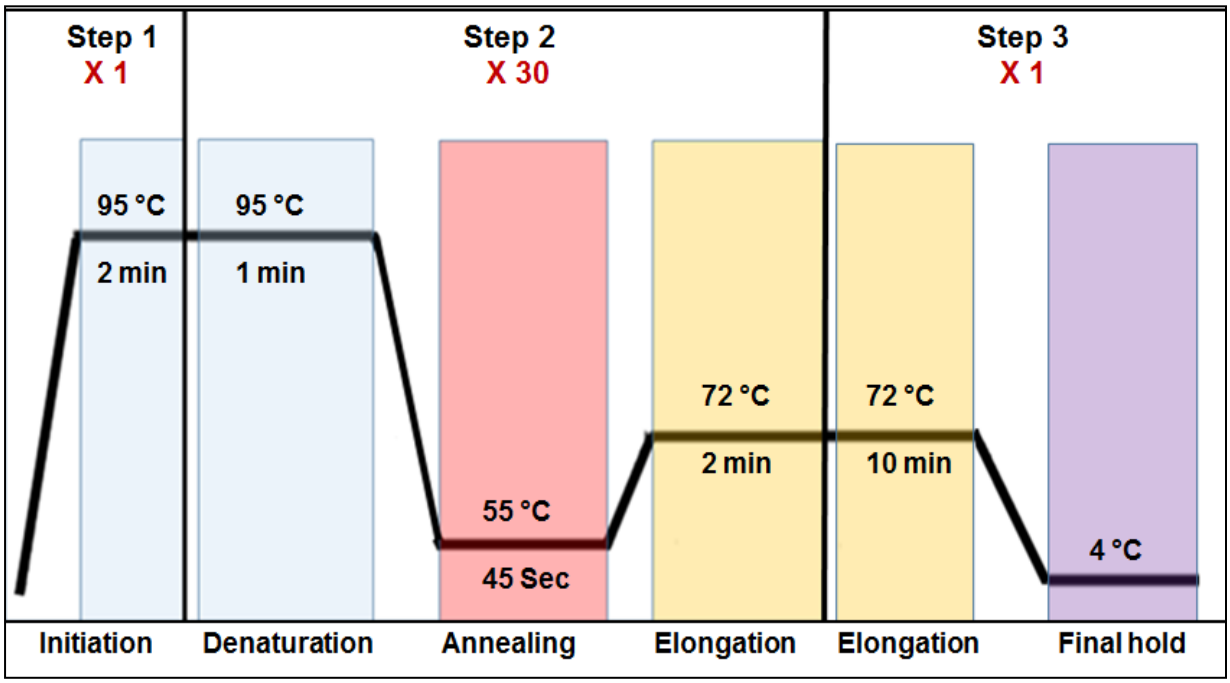


Figure 2.1. **The temperature profile of a typical PCR reaction cycle.**

The PCR reaction consists of three major steps for amplification of target DNA. In the first cycle of the PCR, denaturation was extended to 2min. Second step of reaction consists of 30 cycles. It consists of denaturation of target DNA at 95°C for 2 mins, annealing of primers to the desired target site at 55°C for 45 secs, synthesis of new DNA strand at 72°C for 2 min. In final step, elongation was extended to 10 min before final hold at 4°C.

### 2.2.2 Overview of pGEX-6P-3 vector

Glutathione S-transferase (GST) is a naturally occurring 26 KDa protein which has high affinity for glutathione coupled to a Sepharose matrix (Harper and Speicher, 2011). This is a reversible binding and the protein can be eluted under mild, non-denaturing conditions using reduced glutathione in the elution buffer. pGEX-6P-3 expression vector (GE healthcare Life Science) enables the easy cloning of a variety of protein constructs, transformation of *E. coli*, and selection of transformants by the appropriate antibiotic (ampicillin) (Figure 2.2). A specific PreScission-protease site engineered between the GST moiety and the protein of interest allows removal of the GST from the target recombinant protein. This protease is specific for the core amino acid sequence Leu-Phe-Gln/Gly-Pro and it cleaves between the Gln and the Gly residue, which has been engineered into the pGEX-6P-3 multiple cloning site (MCS). PreScission Protease is a fusion protein which is genetically engineered and consists of human rhinovirus 3C protease and GST. The GST tag permits purification of the recombinant protein by means of Glutathione Sepharose (GS) affinity chromatography. The GST can be easily removed from the sample by re-chromatography on a GS column.

### 2.2.3 Agarose gel electrophoresis

The PCR products were observed for their size by performing Agarose Gel electrophoresis. For preparing a 1% Agarose gel, 0.5 g of Agarose was heated with 50 ml of 1X TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA pH 8.3) to dissolve the agarose. 3 µL of Hydragreen dye (20,000X) (ACTgene) was added after the gel mixture was slightly cooled and then poured it into the gel casting tray containing a 10-well comb. The agarose gel was allowed to set for about 25 min and then comb was removed. The gel was placed in an electrophoresis apparatus and submerged into 1x TAE buffer. 8 µL of the PCR product was mixed with 2 µL of the 6X loading dye (10 mM Tris-HCl (pH 7.6) 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% glycerol 60 mM EDTA) (Fermentas) and loaded into the well. 10 µL of 1 kb DNA ladder (Fermentas) was also loaded to compare the size of the bands. The gel was run at 100V for 45-50 minutes. The DNA was visualized under gel imager at 365 nm (Syngene Imager).

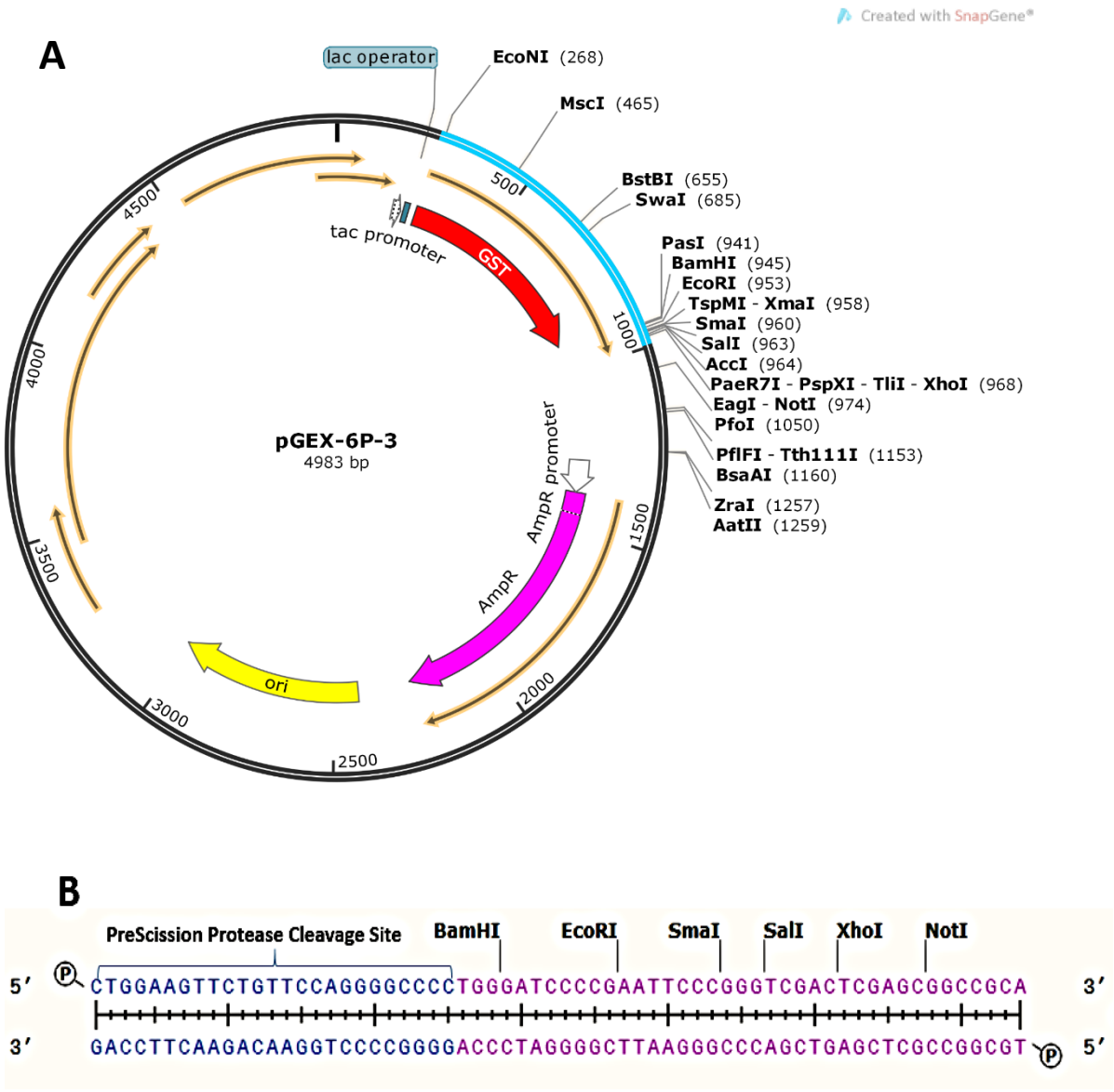


Figure 2.2 . **Overview of pGEX-6P-3 expression system.**

**A.** pGEX-6P-3 (4.9 Kbp) vector showing the Multiple Cloning Site (shown in blue color), the GST gene (shown in red color), the Amp resistance gene (shown in pink color), origin of replication (Ori-shown in yellow color) gene and the endonuclease restriction sites. (The figure was generated using SnapGene software). **B.** Sequence of the multiple cloning site from pGEX-6P-3 vector showing PreScission Protease Cleavage Site and cleavage sites for different restriction enzymes.

### 2.2.4 Ethanol precipitation of PCR product

As a next step to preserve the PCR products from exonuclease activity, ethanol precipitation of PCR products was carried out. 10 % v/v 3M sodium acetate was mixed with the PCR reaction mixture, followed by 3X the reaction volume of 95 % ethanol. The sample was mixed thoroughly and kept in -20°C for 15-20 min. The supernatant was removed carefully after centrifugation 16438 x g at 4°C for 5 minutes. 2x 70 % ethanol was added into the reaction mixture followed by centrifugation at same conditions for 5 minutes. Sodium acetate decreases the pH of the solution to help in precipitation of DNA. Ethanol acts as precipitating agent for DNA, since it is less polar than water and hence it interferes with the DNA-water interaction thus making DNA to precipitate from solution. The pellet was dried using CentriVap Centrifugal Vacuum Concentrator (Labconco) for 15 minutes and then the sample was stored at -20°C. The pellet was dissolved in 20 µL of $H_2O$ before restriction digestion.

### 2.2.5 Restriction digestion and gel extraction of DNA

The product DNA was digested at 37°C O/N with the restriction enzymes *Bam*HI and *Eco*RI (2 µL @20,000 U/mL each), 10 µL of 10X Tango buffer (Fermentas). The same procedure was used to digest the pGEX-6P-3 vector. The digestion of PCR product is required to cut the DNA at appropriate size for amplification of gene of interest. Again ethanol precipitation of the reaction mixture was carried out and the sample was dried by the vacuum concentrator for 15 min. The DNA pellet was re-dissolved in 10 µL of $ddH_2O$. Both the DNA fragments (insert & vector) were further purified by Qiagen Gel Extraction kit. For gel purification, the DNA was loaded on 1 % Agarose gel containing 3 µL of Hydragreen dye. After the DNA extraction step the concentration of the insert and plasmid was calculated by measuring absorbance at 260 nm using Nanodrop Spectrophotometer (Thermo Scientific Nanodrop 2000c).

### 2.2.6 Preparation of competent cells

Competent *E. coli* XL-1 and BL-21 DE3 (BL-21) *E. coli* cells were prepared from the stock culture of respective cells using the following procedure. Cells were first streked onto Luria-Bertani (LB) agar plates without any antibiotic and incubated for 14 - 16 hours at 37°C. On the following day, 5 mL LB broth without antibiotic was inoculated with a single colony and incubated in orbital shaking incubator for 14 - 16 hours at 37°C and 300 rpm. Following incubation, the culture was diluted in a 1:100 proportion with fresh LB broth. The diluted culture

was grown at 37°C and 300 rpm to an optical density (OD 600 nm) of 0.4 to 0.6 and then cells were kept on ice for 5 min. Cells were harvested by centrifugation in a Beckman-Coulter Centrifuge at 2,218 x g and 4 °C for 5 minutes. Supernatant was removed and the cell pellet was gently resuspended into ice cold 0.1 M $MgCl_2$ in a 1:5 v/v ratio. The cell suspension was incubated on ice for 1 hour. Cells were harvested by centrifugation for 5 minutes at 2,218 x g, 4°C. The cell pellet was resuspended in 2 mL of 0.1 M $CaCl_2$ plus 14 % glycerol and incubated on ice for 1 hour. The aliquots were made in eppendorf tubes and flash frozen with liquid $N_2$ and stored at - 80°C.

### 2.2.7 DNA Ligation and transformation in *E. coli* host cells

The ligation of the insert DNA into the plasmid was carried out in a molar ratio of 3:1 (Insert: Vector). 150 ng/µL of insert DNA was mixed with 50 ng/µL of vector (pGEX-6P-3) along with 1 µL of the enzyme T4 DNA Ligase (Fermentas) and 1 µL of 10X T4 buffer (Fermentas). Cloning of the gene of interest in multiple cloning site of pGEX-6P-3, expresses the protein as GST-fusion protein thus allowing to purify the protein by GST affinity purification method. Transformation was done using *E. coli* XL-1 host cells since they have the highest transformation efficiency. 150 µL *E. coli* XL-1 cells were incubated with 5 µL of ligation reaction for 45 minutes on ice and was spread onto pre-warmed LB agar plate containing Amp (final concentration 100 µg/ml). Mock transformation of double digest plasmid was also carried out to check if there is re-ligation of plasmid. The plates were then incubated at 37°C O/N.

### 2.2.8 Recombinant screening

To check for the presence of the insert into the colonies, screening of the insert was performed using PCR. Aliquots of 25 µL of the PCR master mix (10X Thermopol buffer, *Taq* polymerase, 5mM dNTPs, 25 µL of forward (GEX-3) & reverse primer @ 20 µM concentration, 150 µL of $dH_2O$). 25 µL of PCR master mix was inoculated with each colony from the transformation. The standard PCR step was run as per cloning step. Analysis of the PCR reaction was carried out using agarose gel electrophoresis. 5 µL of the sample was run on 1% agarose gel containing 3 µL of the Hydra Green dye. Bands corresponding to insert and plasmid were observed under UV light at 365nm wavelength.

### 2.2.9 Plasmid preparation and purification

For positive colonies, 5 mL of LB media containing 5 µL of 100 mg/mL Ampicillin stock was inoculated and grown overnight at 37°C. Plasmid purification was carried out using QiaGen Plasmid Purification kit, which helps in eluting plasmid DNA. The concentration of plasmid was measured on Nanodrop ($A_{260}$). 10 µL (100 ng/µL) of purified plasmid was digested using restriction enzymes as mentioned in section 2.2.5 and analyzed by Agarose gel electrophoresis.

### 2.2.10 DNA sequencing

The plasmid samples (200 ng) were sent to PBI-NRC at the U of S, Saskatoon, for DNA sequencing to get the sequence of the insert DNA. All the sequence results were compared with genomic DNA (*H. pylori* NCTC 11637) sequence using nucleotide BLAST to check the identity.

## 2.3 Bacterial Strains and growth conditions

*E. coli* strains (Table 2.2) were grown in Luria Bertani (LB) agar plates and LB broth at 37°C incubation. For making 1 L of LB broth, 10 g of Bacto™ Tryptone (Becton, Dickinson and Company, BD), 5 g of Bacto™ Yeast Extract (BD) and 10 g of NaCl was dissolved in 1 L of deionized water. For making LB agar, 15 g of Difco™ Agar was added to 1L LB broth. The broth and LB agar was autoclaved for 20 min at 121°C for sterilization. LB agar was slightly cooled after sterilization and then poured into petri dishes.

For growing *E. coli* cells containing the FliI E193Q mutant plasmid, Hyper Broth™ (Athena Enzyme Systems™) and Glucose Nutrient Mix (Athena Enzyme Systems™) were used. 150 g of Glucose Nutrient Mix was dissolved in 700 mL deionized water and this solution was then filtered using a 0.2 µm filter. Glucose Nutrient Mix was stored at 4°C. To make 1 L of Hyper Broth, 44.5 g of Hyper broth was dissolved in 1L deionized water and autoclaved at 121°C for 15 min. 50 mL of Glucose Nutrient Mix was added to the broth once it is cooled to approximately 50°C.

Table 2.2. **Bacterial Strains used in this project**

| **Bacterial strain** | **Source** |
|---|---|
| *H. pylori* CCUG 17874 - *H. pylori* motile strain (*H. pylori* NCTC 11637) | Culture collection University of Gothenberg, Sweden |
| *E. coli* XL-1 | Stratagene® |
| *E. coli* BL-21(DE3) | Novagen® |

## 2.4 Recombinant Protein methods

### 2.4.1 Protein expression trials

Optimization of protein expression was carried out to find the ideal conditions to get maximum expression of recombinant proteins. To get expression of protein, 40 ng/µL of the plasmid was transformed into *E. coli* BL21 host cells. 1 µL of 40 ng/ µL of plasmid containing the gene of interest was added into 100 µL of competent BL21 cells and kept on ice for 30-40 minutes for transformation. Plating was done on LB agar plates with ampicillin and incubated overnight @ 37°C. Next day 5 mL of LB media with Amp (100 µg/ml) was inoculated with a single colony and incubated O/N at 37°C in Innova 4330 Refrigerated Incubator shaker (New Brunswick Scientific). Next day, 3X 50 mL of LB (50 µL Amp) was inoculated with 500 µL of starter culture and incubated @37°C until first one reach an OD of 0.4; second reach an OD of 0.6 and third reach an $OD_{600}$ of 0.8. Once the desired absorbance was obtained, 0.1 mM of IPTG (Isopropyl thiogalactosidase) for protein expression and incubated for three hours @ 37°C. The cells were pelleted and kept in -20°C.

The same procedure was repeated for expression at 25°C. After addition of IPTG, flasks were incubated at 25°C O/N. Next day, cells were pelleted and kept at -20°C O/N.

Four different samples were prepared for SDS-PAGE analysis.

1. Sample 1(Lysate): - Pellet resuspended in 500 µL 1X PBS buffer and sonicated until moderately clarified. 5 µL of this is removed for SDS PAGE analysis.

2. Sample 2 (Supernatant): - 5 µL supernatant was removed after centrifugation at 15588 x g for 20 minutes.

3. Sample 3 (Pellet): - The pellet was resuspended in 500 µL of PBS and 5 µL was used for SDS-PAGE analysis.

4. Sample 4 (GST-pull down): - 20 µL of 70% GST-sepharose slurry was added to remaining supernatant sample and mixed for 10-20 minutes at room temperature and supernatant was removed by centrifugation. The beads were washed with 200 µL of PBS and resuspended in 20 µL of SDS PAGE loading buffer.

5. 5 µL of loading buffer was added to samples 1, 2 and 3 and boiled for 2 min.

6. The SDS PAGE was run of the samples along with molecular mass markers on 15% SDS gel.

### 2.4.2 Expression of GST fusion proteins and harvesting cells

For transformation of a competent cell line with a desired plasmid, 40 ng/µL of the plasmid was transformed into *E. coli* BL21 host cells & incubated on ice for 30-40 minutes before plating on LB agar plate containing ampicillin. The plate was incubated at 37°C overnight. Next day 50 mL of LB (with 100 µg/ml amp) was inoculated with a single colony from plate and incubated for 37°C overnight in a shaker incubator at 200 rpm. The following day, 2 liters of LB (amp) was inoculated with 2 mL of starter culture and incubated at 37°C till the cells reach an $OD_{600}$ of 0.8. For induction of protein expression, IPTG was added to a final concentration of 0.1 mM (2 mL) and incubated 25°C overnight in a shaker incubator at 200 rpm.

For cell recovery from liquid culture, the cells were spun in a J2-HS centrifuge using the JLA 16.250 rotor at 5393 x g for 30 min at 4°C. The first step in the cell lysis procedure was resuspension of the pelleted cells in ice cold lysis buffer (50 mM Tris-HCl [pH 8.0], 150 mM NaCl, 1 mM EDTA) in a ratio of 1:3 w/v of the pellet to the lysis buffer. The resuspended cells were kept at -20°C until the next step.

### 2.4.3 Lysis of cells

The falcon tubes containing the cell suspension was kept on ice for 5 min and then allowed it to thaw at room temperature. Once the cells completely thaw, 1 mg/mL of Hen egg white lysozyme (Bioshop) was added and incubated on ice for 20 min. The cells were lysed @ 35 psi using a cell disruptor (Constant cell disruptor system) and centrifuged at 20,400 x g in a J2-HS centrifuge using a JA-20 rotor for 60 minute at 4°C. The supernatant was passed through glass

wool and further steps of purification by FPLC (Fast performance liquid chromatography) was performed.

### 2.4.4 Affinity purification of GST-fusion protein

For affinity purification of GST-fusion protein, an ÄKTA FPLC and HR 16/5 column (Amersham Biosciences) were used. HR 16/5 column is packed with 25mL of Glutathoine Sepharose™ 4B beads (Amersham Phramacia Biotech). The column was washed with 2 column volumes of autoclaved $H_2O$ to remove ethanol and then equilibrated with 2 column volumes of 1X PBS (Phosphate buffered saline, 140 mM of NaCl, 2.7 mM of KCl, 10 mM of $Na_2HPO_4$, 1.8 mM of $KH_2PO_4$, pH 7.3). The lysate supernatant was passed through glass wool to remove particulate impurities and then it was applied to the column using a 50 mL Superloop™ (Amersham Pharmacia Biotech) at a flow rate of 1 mL/min. The column was washed with 3 column volume of 1X PBS. The GST fusion protein binds to the column. It was eluted using elution buffer (50 mM Tris HCl pH 8.0, 10 mM reduced glutathione). Proper fractions corresponding to peak of interest were collected and 10 μL of samples were removed for SDS-PAGE analysis.

### 2.4.5 Buffer exchange by dialysis

Buffer exchange by selective diffusion plays an important role in protein purification and helps in removing small, unwanted molecules from the protein. Buffer exchange by dialysis was achieved by using BioDesign Dialysis Tubing™ with a 3,500 molecular weight cut-off (MWCO). The dialysis tube was cut to a required length and soaked in water for 15 min. One end of the tube was tied and sealed with a clip. The protein sample was poured in the tube and other side was also knotted. The sample was subjected to dialysis at 4°C with 4 liter of cleavage buffer (pH 8.0) (50 mM Tris, 150 mM NaCl, 1 mM EDTA) for 16 hours. The buffer was changed for another five hours.

### 2.4.6 Preparation of PreScission protease

PreScission Protease (PSP) (46.4 kDa) is a genetically engineered fusion protein of glutathione S-transferase (GST) and human rhinovirus (HRV) type 14 3C protease (Walker *et al*, 1994). This protease specifically identifies a subset of sequences having a stretch of amino acids, Leu-Phe-Gln/Gly-Pro and cleaves between the Gln and Gly residues (Cordingley *et al*, 1990). After cleavage of GST, it can be easily removed by using Glutathione Sepharose™ 4B.

Expression and purification of GST-PSP was done using the same method as explained in section 2.5. The lysis buffer (Buffer P) used for GST-PSP purification was prepared by dissolving 50 mM Tris-HCl, 150 mM NaCl and 10 mM EDTA. This buffer was autoclaved and 10 mM β-ME (β-mercapto-ethanol) was added just before resuspending cells.

Resuspended cell solution was kept at -20°C for 24 hours and then lysis and purification was carried out using Glutathione sepharose 4B$^{TM}$ column as explained in section 2.4.4. For purification of GST-PSP, 1X buffer P was used instead of 1X PBS. After running 12% SDS gel, appropriate GST-PSP fractions were collected and dialyzed against 4 liter of cold buffer P for 24 hours. Fractions with pure GST-PSP were pooled and then concentrated till 15 mg/mL using an Amicon® Ultra-15 millipore filter 10,000 MWCO. The concentration of GST-PSP was determined by Nanodrop ($A_{280}$). 25 % glycerol stock of the protein was prepared in small aliquots and kept in -80°C after freezing them in liquid $N_2$.

### 2.4.7 GST cleavage using PreScission protease and removal of GST tag

The protein solution was transferred to a falcon tube after dialysis. The GST tag was removed using PreScission protease. The protein was mixed with 80 units of PreScission protease to cleave the GST tag and 4 mM DTT was also added. DTT reduces the disulfide bonds of proteins and thus helps to prevent cysteine residues from forming intramolecular and intermolecular disulfide bonds. The protein solution was kept at 4°C for a minimum of 16 hours. The GST cleaved protein was then subjected to a 2$^{nd}$ round of GST purification wherein free GST binds to the Glutathione Sepharose$^{TM}$ 4B beads and pure protein comes out in flow through fractions.

### 2.4.8 Ion exchange chromatography

Ion exchange chromatography is an effective method used for protein purification. In anion exchange chromatography, negatively charged protein moieties are attracted to a positively charged solid support and then increasing gradient of salt is applied to elute the desired protein. The flow through fractions from 2$^{nd}$ step of GS affinity purification were collected, pooled and subjected to dialysis in buffer A (50 mM Tris pH 8.0, 25 mM NaCl) for ion-exchange chromatography. The protein was purified using an anion exchange column (Source$^{TM}$ 15 Q column) packed in a HR 16/5 column. The column was pre-equilibrated with 50 mL of buffer A. The protein sample was injected in the column using a superloop at 1 mL/min flow rate. The

protein was eluted with a linear gradient of 25-500 mM NaCl over three column volume using elution buffer B (50 mM Tris-HCl pH 8.0, 2 M NaCl) and appropriate fractions containing protein of interest were collected.

### 2.4.9 Concentration of recombinant proteins

The concentration of protein was determined from $A_{280}$ (absorbance at 280 nm) by Nanodrop Spectrophotometer (Thermo Scientific Nanodrop 2000c). $A_{280}$ was based on extinction coefficient of Tyr and Trp residues present in the protein (1490, 5500 respectively). The extinction coefficient of all the proteins was determined by ExPASy ProtParam tool and then the concentration was determined by dividing the absorbance by the extinction coefficient. The protein was centrifuged using an Amicon® Ultra-15 centrifugal filter device, 3 kDa Molecular Weight Cut Off (MWCO) to a final concentration of 10 mg/mL. 20 mM Bis-Tris propane pH 7.0, 100 mM NaCl, 2 mM TCEP was used as a buffer. Buffer exchange helps in removing contaminating charged ions in protein solution. The buffer pH affects the solubility of a protein. Finally, the concentration of protein was measured and different concentrations of proteins were run on 15% SDS gel to determine the purity.

### 2.4.10 Size exclusion chromatography (SEC)

Gel filtration or size exclusion chromatography separates molecules/ proteins based on their molecular size. SEC was performed on a Superdex$^{TM}$ 75 or a Superdex$^{TM}$ 200 column depending on molecular size of proteins. The two columns differ in the average pore size of the matrix hence help in better separation of proteins depending on the molecular size of proteins. Superdex$^{TM}$ 75 column is generally used for proteins with an approximate size between 70 kDa and 2 kDa. Superdex$^{TM}$ 200 column is generally used for proteins with an approximate size between 400 kDa and 7 kDa. The column was washed with dd$H_2O$ and then equilibrated with 20 mM Tris pH 8.0 and 100 mM NaCl. The protein was injected to the column using either a 500 µL or a 2 mL sample loop (GE Life Science). The flow rate used for sample injection and elution was 0.5 mL/min. The max pressure for both the columns is 1 MPa. After elution, the column was cleaned by 3 CV of $H_2O$ and 3 CV of 20 % ethanol, and stored at room temperature. Based on elution volume, molecular mass of the protein was calculated using a standard calibration curve.

### 2.4.11 Site Directed Mutagenesis

Mutagenesis of FliI was performed using QuikChange® Site-Directed Mutagenesis Kit (Agilent technologies).

**Mutant strand synthesis reaction (Thermal Cycling)**: - A series of sample reactions were set up by mixing 10ng, 20ng and 50ng of the plasmid DNA with 5 µL of 10X reaction mixture (100 mM KCl, 100 mM 100 mM$(NH_4)_2SO_4$, 200 mM Tris-HCl (pH 8.8), 20 mM MgSO4, 1% Triton® X-100, 1 mg/ml nuclease-free BSA). 1.25 µL of each primer (125 ng) (Primers ML097 and ML098; See Table 2.2 for primers) was added along with 1 µL of dNTP mix. Finally, 1 µL of *PfuTurbo*™ DNA polymerase was added (2.5 U/ µL). Thermal cycling of the PCR was carried out in a Bio-Rad MyCycler. The cycling parameters for the mutagenesis are listed in Table 2.3.

Table 2.3. **Cycling parameters for the mutagenic DNA strand synthesis by QuikChange® Site-Directed Mutagenesis Kit.**

| Segment | Number of cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 95°C | 30 seconds |
| 2 | 16 | 95°C | 30 seconds |
| | | 55°C | 1 minute |
| | | 68°C | 5 minute |

Following thermal cycling, the reaction was placed on ice for 2 minutes to cool the reaction around 37°C.

*Dpn*I digestion of the amplification products: - The products were digested with 1 µL of *Dpn*I (10 U/µL stock) enzyme for 1 hour at 37°C which digests the methylated, non-mutated parental plasmid.

**Transformation of XL1-Blue Supercompetent cells** - 1 µL of the *Dpn*I treated DNA was added to previously thawed 50 µL XL-1 Supercompetent cells (Agilent technologies) and incubated on ice for 30 minutes. The reaction was heated at 42°C for 45 seconds then placed on ice for 2 minutes. The reaction was incubated at 37°C for 1hour in a shaker incubator at 200 rpm after addition of 0.5 mL of NZY+ broth. For preparing $NZY^+$ broth, 2 g of NZ amine (Casein

hydrolysate) (Sigma-Aldrich), 1 g yeast, 1 g NaCl was dissolved in 200 mL of deionized water and autoclaved for 20 min at 121°C. 2.5 mL of 1M $MgCl_2$, 2.5 mL of 1M $MgSO_4$ and 4 mL of 20% glucose was added to above media prior to use.

The 250 µL of reaction volume was plated on LB plate supplemented with 100 mg/mL of ampicillin. The plates were incubated at 37°C for 16 hours. The following day, 5 mL of LB media with ampicillin was inoculated with single colony and incubated at 37°C for 16 hours. Plasmid preparation was carried out as mentioned previously in section 2.3.9 and the concentration of mutagenic plasmid was determined by Nanodrop ($A_{280}$ = (0.1 %) = 0.987). 10 µL of the plasmid was digested to check the size of the inserted gene. The mutagenized plasmid (200 ng) was sent for sequencing PBI-NRC at U of S as mentioned in section 2.3.10.

## 2.5 Protein Analysis

### 2.5.1 SDS-PAGE

For separation and visualization of proteins, Sodium-dodecyl sulfate Polyacrylamide gel electrophoresis (SDS-PAGE) using Laemmli method was performed (Laemmli, 1970). All SDS-PAGE gels were cast with 15% or 12% resolving gel and 4 % stacking gel by using Hoefer SE245 Mighty Small Dual Gel Caster (GE Life Science). For making 15 % resolving gel, 1.85 mL 40 % acrylamide, 1.38 mL 1.5 M Tris-HCl [pH 8.8 at 25°C], 50 µL 10 % (v/v) SDS, 25 µL 10 % APS, 5 µL TEMED were mixed and volume was made up to 5 mL with $H_2O$. For making 4 % stacking gel, (0.25 mL 40 % acrylamide, 0.62 mL 0.5 M Tris-HCl [pH 6.8 at 25°C], 25 µL 10 % (w/v) SDS, 25 µL 10 % APS, 5 µL TEMED were mixed and volume was made up to 2.5 mL with $H_2O$). A 12% resolving gel was made by adjusting the amount of 40 % acrylamide and $H_2O$.

5 µL of samples were mixed with 5 µL of sample loading buffer (0.5 M Tris-HCl [pH 6.8], 10 % glycerol, 10 % SDS, 5 % (w/v) β mercaptoethanol, 1 % (w/v) Bromophenol blue). The sample was then boiled for 2 min, and 10 µL was loaded on the gel. SDS-PAGE was run at a 115 V until the dye front reaches the bottom of the gel (approximately 45 minutes).

The running buffer for SDS gel consisted of 1X SDS buffer (25 mM Tris-HCl (pH 8.2), 200 mM glycine, 0.1 % (w/v) SDS) diluted from a 10X stock solution. The SDS gel staining buffer consisted of Coomassie staining solution (40 % ethanol, 10 % v/v glacial acetic acid, 0.1 % w/v Coomassie Brilliant Blue). The gel was stained with Coomassie staining buffer for 30

min by gently mixing on an IKA® VIBRAX VXR Basic Orbital Shaker. Destaining solution was prepared by mixing 7.5 % v/v glacial acetic acid, and 10 % v/v ethanol. The gel was kept for destaining in destaining solution for 4-5 hours at room temperature. Photographic images of the SDS-gel were taken using Gel Imager (Syngene Gel Imager).

### 2.5.2 Determination of protein concentration

$A_{280}$ of proteins was determined using Nanodrop spectrophotometer (Thermo Scientific Nanodrop 2000c). The spectrophotometer was blanked using 2 µL $ddH_2O$ and then $A_{280}$ of 2 µL of protein was determined. The concentration of protein was then calculated using theoretical absorbance of 1 g/L of protein at 280 nm (obtained from ProtParam).

### 2.5.3 Protein – protein interaction studies-GST-pull-down assay

GST pull-down assay is an *in vitro* technique used to check protein-protein interaction (Brymora *et al*, 2004). GST pulldown experiment consists of a GST-fusion protein (the bait) mixed with an excess of the recombinant tag-free protein (the prey). The bait protein bound to Glutathione Sepharose™ 4B resin was incubated with the prey protein and then extensively washed with buffer. If there is interaction between the prey and the bait protein, it can be observed by the retention of prey protein on a gel containing the glutathione sepharose resin. 60 µg of the GST-fusion protein was added to approximately 20 µL of sepharose beads (70 % solution) and incubated for 1 hour at 4°C and gently mixed on a Speci-Mix Aliquot Mixer (Barnstead International). Beads were pelleted by centrifugation using a Microfuge® 18 Centrifuge (Beckman Coulter™) at 1476 x g for 1 min. The supernatant was discarded and pellet was washed four times with 400 µL of 1X PBS pH 7.3. As a next step, the purified protein was added and incubated for two hours at 4°C on a Speci-Mix Aliquot Mixer (Barnstead International). Again the beads were washed with 1X PBS pH 7.3, after centrifugation at 1476 x g for 1min. The supernatant was discarded and the remaining pellet was mixed with 15 µL of sample loading buffer boiled for 2 min and then analyzed by SDS-PAGE.

## 2.6 Protein crystallization

When a homogenous protein solution reaches its limiting concentration, the protein no longer remains in the same state of solubility. This phenomenon forms the basis of protein crystallization experiments (Asherie, 2004). By changing the experimental conditions, the crystallographer tries to exceed the solubility limit of the protein so that it results in production

of crystals. For the crystallization experiment, a phase diagram shows the state of protein as a function of various crystallization parameters (Chayen, 2004). The phase diagram (Fig. 2.3) is used to quantify the effect of various parameters (protein concentration, precipitant concentration etc.) on the production of crystals. Thus, phase diagrams give basic criteria to design the experiment for protein crystallization and optimization. A typical crystallization phase diagram, consists of four zones representing different degrees of supersaturation. The precipitation zone- an area of high supersaturation; moderate supersaturation zone - spontaneous nucleation takes place, the metastable zone - crystals are stable and may grow; and a zone of under-saturation - the protein is completely dissolved and chances of crystallization are less.

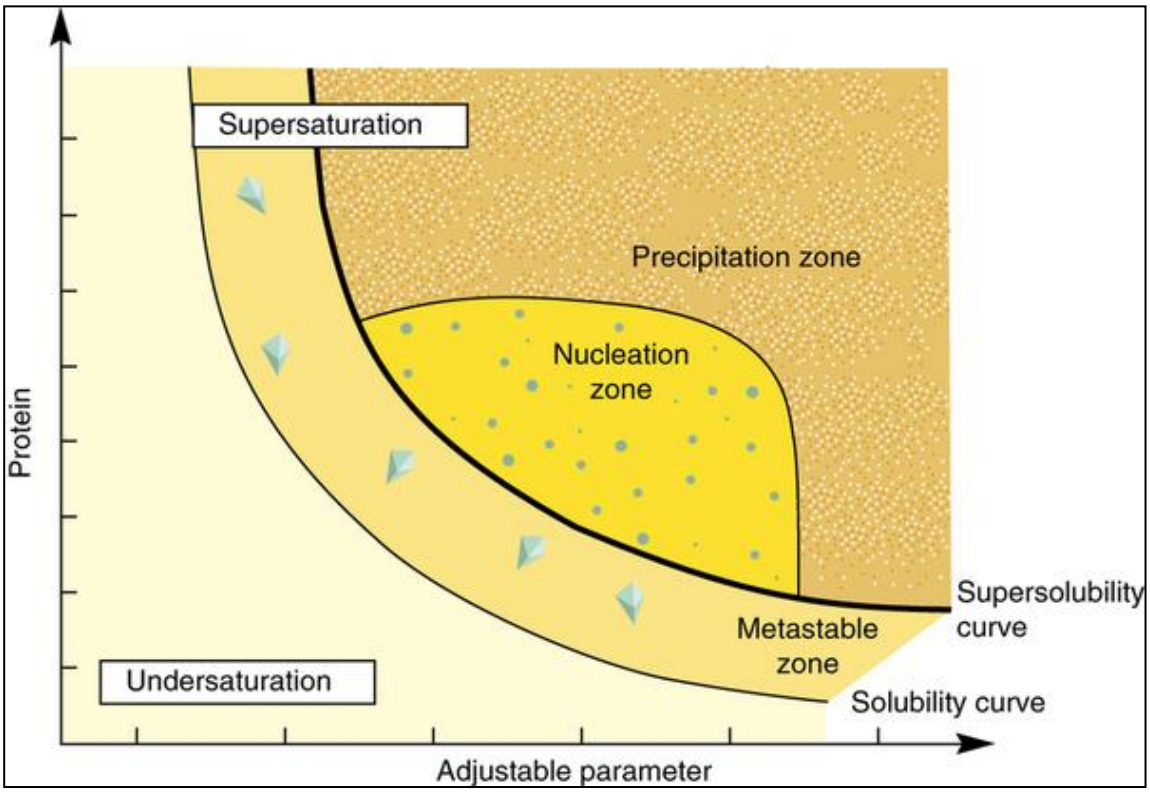


Figure 2.3. **Phase diagram for crystallization.**

Phase diagram showing various stages of crystallization experiment. An area of high supersaturation where protein precipitates; an area of moderate supersaturation in which nucleation of protein is observed; a metastable zone where crystals can be observed and growth of crystals can be achieved due to lower supersaturation; an area of under saturation where protein is in fully soluble state and crystals are not observed. (Figure adapted from Chayen, 2004)

### 2.6.1 Hanging drop vapor diffusion method

Protein crystallization experiments were performed using a hanging drop vapor diffusion technique (Chayen, 1998). The vapor diffusion technique is a commonly used method for crystallization. It is based on the presence of precipitant which absorbs water from the crystallization drop hence driving the solution into supersaturated state and favoring crystallization. 8-10 mg/mL of the protein was used for setting up the crystallization drops in Linbro plates (Hampton Research). The screens were set up using the Gryphon Robot (Art Robins Instruments). The individual hanging drops were equilibrated against 1 mL of reservoir solution. Each drop consists of 1 μL of the purified protein solution mixed with 1 μL of the reservoir solution.

For optimization of crystallization conditions various parameters like initial protein concentration, incubation temperature, the effective pH of the reservoir/protein droplet, type of precipitant, additives such as salts, and the type of buffer were used to get favorable conditions for crystal growth.

### 2.6.2 Seeding

Seeding is a powerful technique used for the separation of nucleation and growth of crystals. In this method, already grown crystals are used as seeds and introduced into new drops. These drops are equilibrated at lower levels of supersaturation prior to seeding. Seeding in conjunction with other parameters such as the precipitant and protein concentrations, pH and additives can be optimized to get better quality crystals (Bergfors, 2003).

## 2.7 Protein structure determination by Molecular Replacement and model building.

For X-ray diffraction and data collection, the protein crystals were taken to Canadian Light Source (U of S, Saskatoon). The data was collected at CMCF 08ID-1 beamline or at CMCF 08Ba-1 BM beamline. The diffraction data was processed with HKL2000 (Otwinowski and Minor, 1997). For solving the structure of full length FliS, the PHASER molecular replacement method was used (McCoy *et al*., 2007). Molecular Replacement is the method used to derive the phase information and solve a phase problem using the previously determined model of the structure. Molecular Replacement (MR) provides an initial starting model for refinement. In the absence of phases (Φ), it is difficult to complete the Fourier transform relating the experimental data in reciprocal space to real-space electron density, which in turn is used to build the atomic

model. MR tries to find the model which best fits experimental intensities among already known structures. The model was built using COOT (Emsley *et al*., 2010). The first step of model building was to add backbone residues as well as side chains to the unoccupied areas of the electron density map and then the model was run thorugh the refinement program. The R values (R-free and R-work) were used to monitor the progress of refinement process. The process of model building and refinenment was repeated until the R values become stable and did not decrease any further.

## 3. Results

### 3.1 Characterization of *H. pylori* FliH

#### 3.1.1 Cloning of *H. pylori* FliH (2-258)

For the cloning of the two new truncated variants of FliH it was necessary to clone the full length FliH (2-258) as a template. For cloning of *H. pylori FliH* (2-258), *H. pylori* NCTC 11637 genomic DNA was used as a template. Standard cloning protocol was used as mentioned in materials and methods. The full length FliH (2-258) comprised of 764 bp or 257 amino acids. After plasmid purification it was necessary to digest the plasmid using *Bam*HI and *Eco*RI to check the size of gene of interest and vector. Proper insertion of the gene of interest requires verification of correct orientation, absence of unwanted mutations, and ability to express the recombinant protein in-frame. The desired gene fragment of interest was verified by plasmid digestion (Figure 3.1A) as well as by DNA sequencing (sequencing results not shown). Concentration of pGex-6P-3-FliH (2-258) plasmid after successful plasmid preparation was found to be 63.5 ng/µL.

#### 3.1.2 Cloning of two truncated FliH variants: FliH (57-258) & FliH (73-258)

Previous characterization studies of full length *H. pylori* FliH in Dr. Moore's lab did not yield any crystals. Based on the sequence conservation of *H. pylori* FliH, previously FliH 89-258 was cloned, purified and crystallized in Dr. Moore's lab but due to poor diffraction quality of the crystals, the structure of the protein could not be determined. Increasing the overall size of the construct may help in improving stability by means of affecting the subsequent secondary structure of the protein or by enhancing crystal lattice contacts between FliH molecules. Hence longer constructs were designed which are listed as follows: FliH (57-258) & FliH (73-258). The cloning was done for both the FliH constructs using the protocol mentioned in section 2.2. The restriction digested plasmids were run on agarose gel (Figure 3.1B and 3.1C). Based on the sequence (not shown) and size of DNA fragments on agarose gel it is confirmed that cloning of both the constructs was done successfully. The plasmid concentration of pGex-6P-3-FliH (57-258) and pGex-6P-3-FliH (73-258) was found to be 55.55 ng/µL and 152.2 ng/ µL.

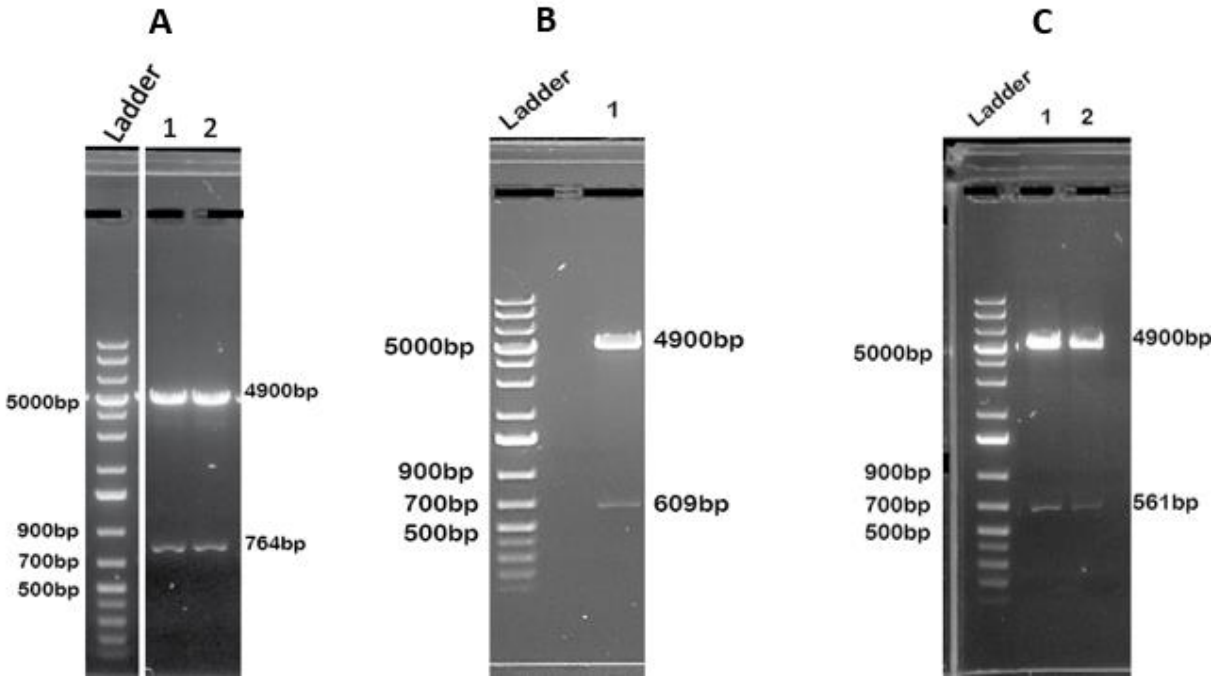


Figure 3.1. **Agarose Gel Electrophoresis of FliH (2-258), FliH (57-258) and FliH (73-258) plasmid digests.**

10 µL of plasmid was digested at 37°C O/N with the restriction enzymes *BamHI* and *EcoRI* (0.5 µL @20,000 U/mL each), 5 µL of 10X Tango buffer (Fermentas). 8 µL of the sample was mixed with 2 µL of the 6X loading dye (Fermentas) and loaded on 1% agarose gel containing 3 µL of Hydragreen dye. 10 µL of 1 kb DNA ladder (Fermentas) was loaded to compare the size of the bands. The gel was run at 100V for 45-50 minutes. The DNA was visualized under gel imager at 365 nm (Syngene Imager). A. Sample 1 & 2 shows the digested FliH (2-258) at 764 bp and pGex-6P-3 plasmid at 4900bp. B. The digested FliH (57-258) at 609 bp and pGex-6P-3 plasmid at 4900bp. C. Sample 1 & 2 shows the digested FliH (73-258) at 561 bp and pGex-6P-3 plasmid at 4900bp.

### 3.1.3 Optimization of protein expression FliH (57-258) & FliH (73-258)

Expression and purification of both the FliH fragments was done using *E. coli* BL21cells which provides an easy and fast system for overexpression of proteins. The initial screening for maximum protein expression conditions was carried out using small-scale batch preps; harvested cells were lysed by sonication and separated into soluble and insoluble fractions by centrifugation. The resulting supernatant was subjected to small-scale GST pull-down as mentioned in materials and methods section 2.5.3. The GST pull-downs were analyzed by 15% SDS PAGE gels.

For determination of optimal conditions for expression of proteins different induction temperatures (25°C and 37°C) and different optical densities at 600 nm wavelength ($OD_{600}$) (0.4, 0.6, and 0.8) were used. The small scale protein expression trials for proteins, FliH (57-258) & FliH (73-258) (Figure 3.2 and Figure 3.3 respectively) shows the difference in protein solubility at different temperature and ODs. A significant proportion of the over-expressed protein was found to be soluble at 25°C and 0.8 OD, enabling further large-scale expression and purification of both the FliH constructs.

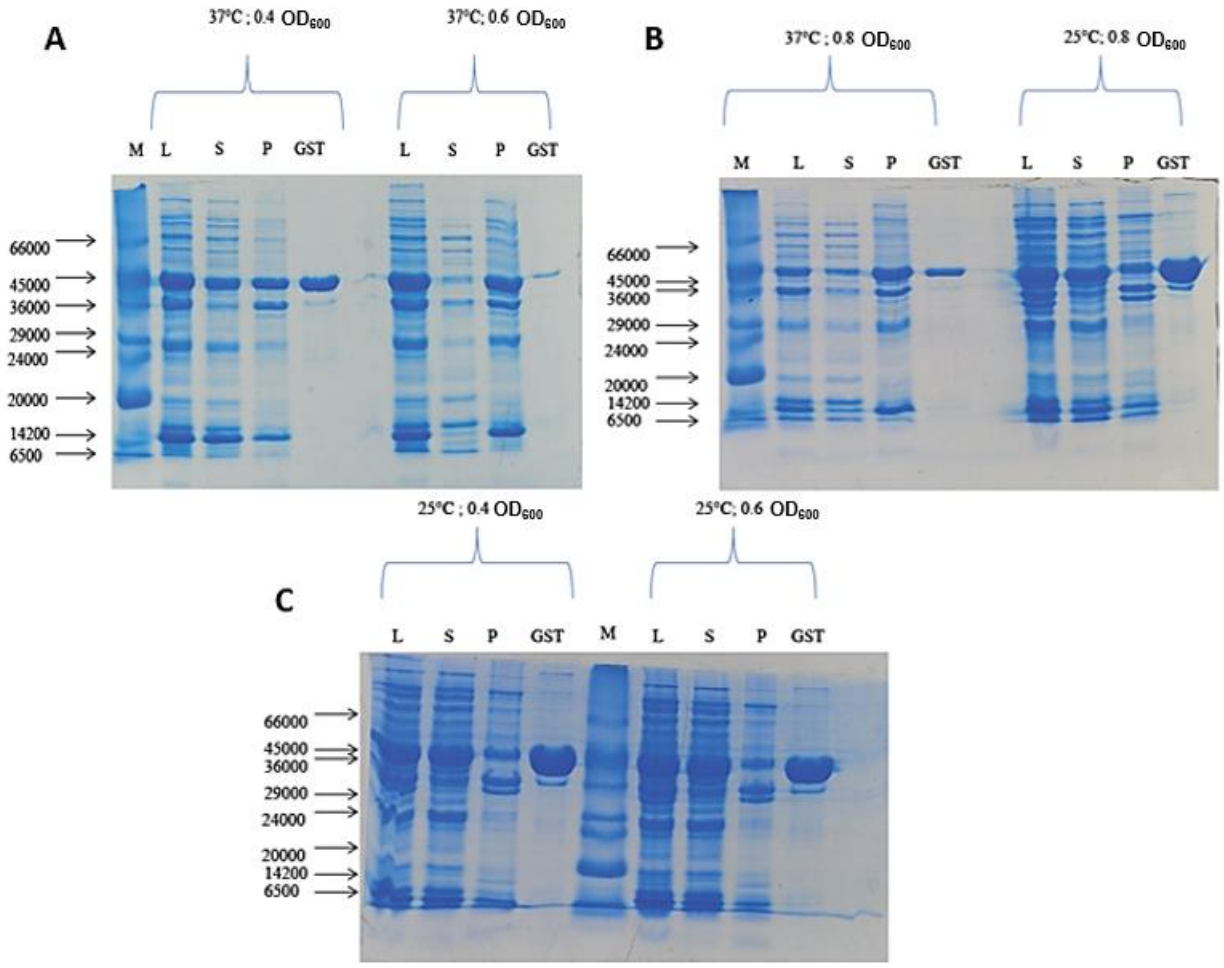


Figure 3.2. **Optimization of expression for FliH (57-258).**

FliH (57-258) was expressed in *E. coli* BL21 cells grown at 37°C or 25°C till desired $OD_{600}$ and then protein expression was induced by adding 0.1 mM IPTG. Different samples were incubated at different temperatures. 5 µL of protein samples were analyzed by 15% SDS-PAGE. (A) Protein expression at 37°C at two different OD 0.4 and 0.6. (B) Protein expression at 37°C at 0.8 $OD_{600}$ and 25°C at $OD_{600}$ 0.8. (C) Protein expression at 25°C at $OD_{600}$ 0.4 and 0.6. M- Marker, L- Lysate, S- Supernatant, P- Pellet, GST- GST pull down.

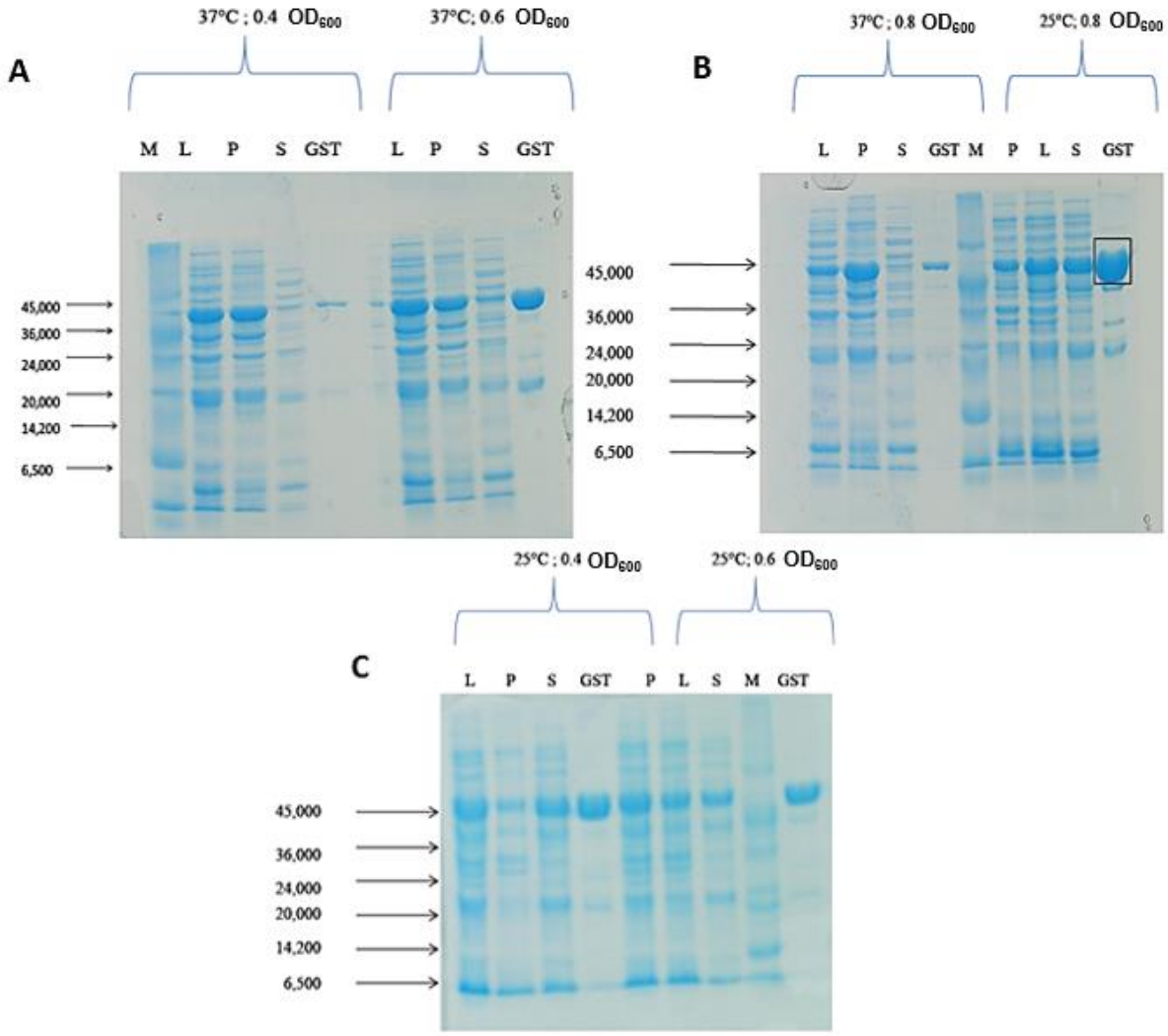


Figure 3 3. **Optimization of expression for FliH (73-258).**

FliH (73-258) was expressed in *E. coli* BL21 cells grown at 37°C or 25°C till desired $OD_{600}$ was reached and then protein expression was induced by adding 0.1 mM IPTG. Different samples were incubated at different temperatures. 5 µL of protein samples were analyzed by 15% SDS-PAGE. (A) Protein expression at 37°C at two different OD 0.4 and 0.6. (B) Protein expression at 37°C, 0.8 $OD_{600}$ and at 25°C, $OD_{600}$ 0.8. (C) Protein expression at 25°C $OD_{600}$ 0.4 and 0.6. M- Marker, L- Lysate, S- Supernatant, P- Pellet, GST- GST pulldown of GST-FliH fusion protein.

### 3.1.4 Protein expression and GST affinity purification of FliH (57-258) & FliH (73-258)

Following optimization of protein expression, large scale (4 L) expression of both the proteins, (FliH 57-258 and FliH 73-258) was carried out. Purification was done by affinity chromatography as outlined in the section 2.5.2. The same GS-affinity purification technique was used to purify all the proteins mentioned in this study. The cells were lysed using a cell disruptor, and centrifuged at 20,400 x g using a JA-20 rotor for 60 min at 4°C to separate the soluble fraction from the insoluble fraction. The lysate supernatant containing the soluble GST-FliH (57-258) or GST-FliH (73-258) protein was purified by FPLC GS-affinity chromatography. The molecular weights of recombinant GST-FliH (57-258) and GST-FliH (73-258) were determined by ExPASy ProtParam tool (49 kDa and 44 kDa respectively). The correct sized GST-fusion proteins were over-expressed and verified by SDS-PAGE, as the protein band migrated at its predicted molecular weight. The chromatogram of the first GST affinity purification step verified the desired interaction of the GST tagged protein with the stationary phase (glutathione sepharose), which enabled purification from the unwanted whole cell extract. The FliH-GST fusion proteins were eluted from the glutathione sepharose column using reduced glutathione in elution buffer, this step is referred to as 1$^{st}$ GST in this study. The GST tag was then cleaved by the addition of PreScission protease which allows removal of the GST-tag after second passage over the glutathione sepharose column, this step is referred to as 2$^{nd}$ GST. Small samples of the proteins were removed after every step of purification for analysis by SDS-PAGE.

**Affinity purification of FliH (57-258)**

Figure 3.4 and Figure 3.5 shows the SDS gel images of FliH (57-258) fractions after each step of affinity purification. The *E. coli* lysate supernatant was injected onto GS column. The GST-FliH (57-258) fusion protein binds to the column and flow-through was washed away with 1X PBS. The bound fusion protein was then eluted using 50 mL elution buffer (50 mM Tris HCl pH 8.0, 10 mM reduced glutathione). Figure 3.4 shows the fractions analyzed by SDS PAGE. The GST-FliH (57-258) has a predicted molecular weight of approximately 49 kDa. The predicted molecular weight FliH (57-258) from its sequence is approximately 23 kDa. After dialysis of the eluted fusion protein in the cleavage buffer (pH 8.0) (50 mM Tris, 150 mM NaCl, 1 mM EDTA), 80 units of PreScission Protease and 4 mM DTT was added. The protein solution was kept at 4°C for a minimum of 16 hours. The GST cleaved FliH (57-258) was then subjected to a 2$^{nd}$

round of GST purification wherein free GST binds to the Glutathione Sepharose$^{TM}$ 4B beads and pure protein comes out in flow through fractions (Figure 3.5).

FliH (57-258) have predicted isoelectric point of 5.05, therefore it was decided to further purify it by anion exchange chromatography using a 10 mL Source-Q (SQ) column. The column was equilibrated using 50 mL of buffer A (50 mM Tris-HCl pH 8.0, 25 mM NaCl). Prior to injection, protein sample was subjected to dialysis in buffer A for about 16 hrs at 4 °C. The protein was then injected on SQ column at a flow rate of 1 mL/min. Figure 3.6 shows that FliH (57-258) binds to the column and it was eluted by applying a linear gradient of 25-500 mM NaCl over three column volumes using buffer A and elution buffer B (50 mM Tris-HCl pH 8.0, 2 M NaCl). Figure 3.6B shows that the protein was purified to homogeneity.

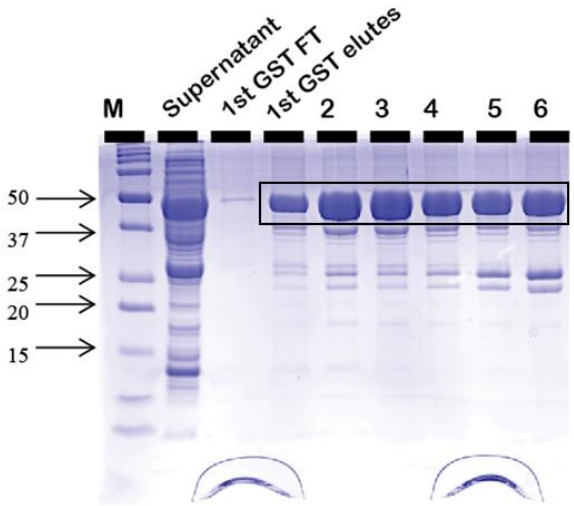


Figure 3.4. **Purification of FliH (57-258).**

Purity of the FliH (57-258) at various stages of purification was analyzed by 15% SDS-PAGE. 5 µL of each sample was boiled with 5 µL of loading buffer and loaded on 15% SDS gel. SDS-PAGE was allowed run at a 115 V until the dye front reaches the bottom of the gel. The gel was stained with Coomassie staining buffer for 30 min and then destained in destaining solution (7.5% v/v glacial acetic acid, and 10% v/v ethanol) for 4-5 hours. Photographic image of the SDS-gel was taken using Gel Imager (Syngene Gel Imager). M- Marker, Supernatant, *E. coli* lysate supernatant; 1$^{st}$ GST FT, flow-through fraction obtained after applying lysate supernatant onto the glutathione sepharose column;1$^{st}$ GST elutes (2-6), eluted fractions obtained after applying lysate supernatant onto the glutathione sepharose column.

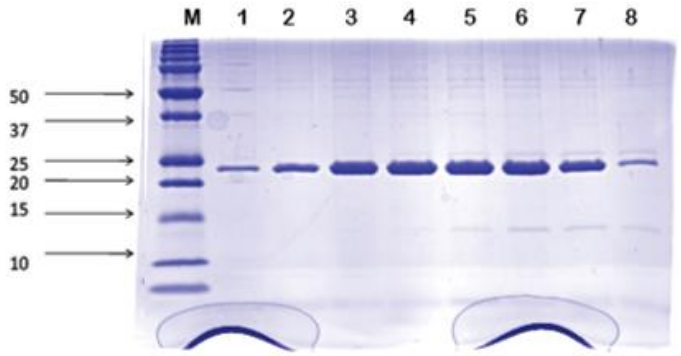


Figure 3.5. **Purification of FliH (57-258).**

2nd GST flow-through fractions analysed by 15 % SDS-PAGE showing purity of the FliH (57-258) after cleavage of GST tag. FliH (57-258) migrates with the molcular weight of 23 kDa on 15 % SDS gel.

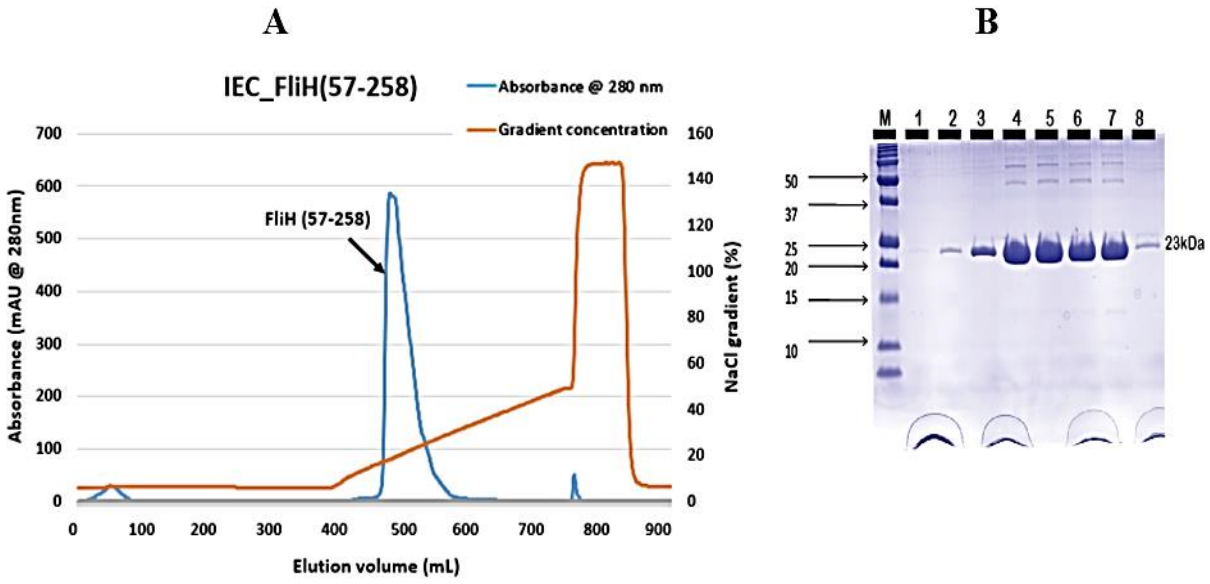


Figure 3.6. **Purification of FliH (57-258) using Source-Q Anion Exchange Chromatography.**

A. Elution profile for purification of FliH (57-258). The column was equilibrated in 50 mL of buffer A (50 mM Tris-HCl pH 8.0, 25 mM NaCl); the flow rate of injection was 1 mL/min. The bed volume of the column was 10 mL. Protein elution was monitored by absorbance (mAu) at 280 nm. B. SDS-PAGE of FliH (57-258) anion exchange chromatography. SQ elute peak fractions were analysed by 15 % SDS gel. M-Protein marker; Lanes (1-8) - Elute peak fractions.

Following anion exchange chromatography, FliH (57-258) was concentrated as mentioned in section 2.4.9.

**Affinity purification of FliH (73-258)**

FliH (73-258) was purified in the same manner as FliH (57-258). Figure 3.7 shows the SDS gel of the fractions after successive stages of purification. FliH (73-258) has molecular weight of 21 kDa and predicted molecular weight of GST- FliH (73-258) fusion protein is approximately 47 kDa. The fusion protein remains bound to GS column which was then was then eluted using 50 mL elution buffer (50 mM Tris HCl pH 8.0, 10 mM reduced glutathione). Figure 3.7 shows the fractions analyzed by SDS PAGE. The GST tag was cleaved similar to explained earlier in this study. The sample was subjected to 2nd round of purification by injecting onto GS column. The FliH (73-258) comes out as flow-through (Figure 3.7).

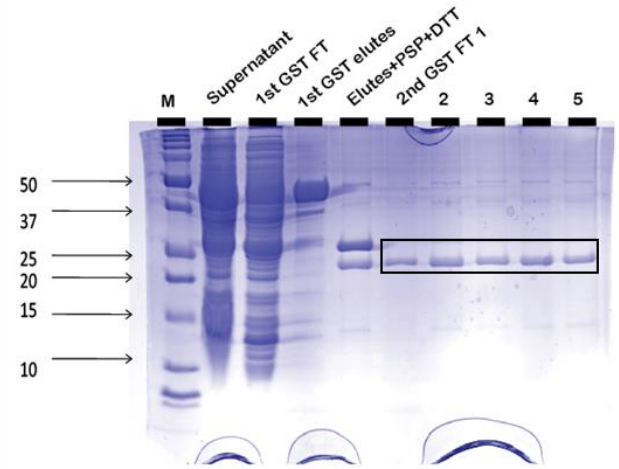


Figure 3.7. **Purification of FliH (73-258).**

Purity of FliH (73-258) at various stages of purification was analyzed by 15 % SDS-PAGE. M-Marker; Supernatant - E. coli lysate supernatant; 1st GST FT - flow-through fraction obtained after applying lysate supernatant onto the glutathione sepharose column; 1st GST elutes- The eluted GST - FliH (73-258) (47 kDa) fractions obtained after applying lysate supernatant onto the glutathione sepharose column; Elutes+PSP+DTT - After dialysis of 1st GST elutes, 80 units of Prescision protease and 4 mM DTT was added to the protein sample for cleavage of GST tag; 2nd GST FT (1-5) - flow-through fractions from the 2nd GST purification FliH (73-258) (21 kDa).

FliH (57-258) have predicted isoelectric point of 5.23. Figure 3.8 shows the typical ion exchange elution profile for this protein. All the fractions from the peak were analyzed by running a 15% SDS gel. Figure 3.8 shows the fractions after ion exchange chromatography. From the gel it can be stated that this protein was purified to homogeneity and can be used for crystallization. All the fractions containing protein of interest were pooled and concentrated using 20 mM Bis-Tris propane pH 7.0, 100 mM NaCl, 2 mM TCEP. Different concentrations of protein were analyzed by SDS PAGE (Figure 3.8). The protein was concentrated as mentioned in methods section to 10 mg/mL using an Amicon® Ultra-15 centrifugal filter device, 3 kDa Molecular Weight Cut Off (MWCO) and used for crystallization studies.

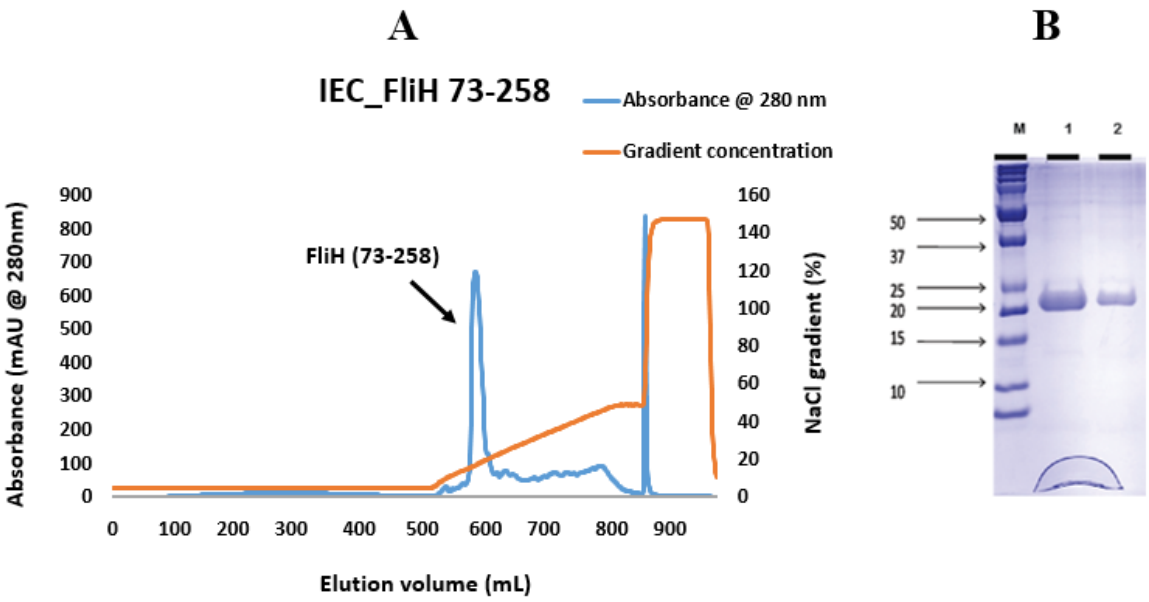


Figure 3.8. **Purification of FliH (73-258) using Source-Q Anion Exchange Chromatography.**

A. Elution profile for purification of FliH (73-258). The column was equilibrated in 50 mL of buffer A (50 mM Tris-HCl pH 8.0, 25 mM NaCl); the flow rate of injection was 1 mL/min. The bed volume of the column was 10 mL. Protein elution was monitored by absorbance (mAu) at 280 nm. B. SDS-PAGE of FliH (73-258) anion exchange chromatography. SQ elute peak fractions were analysed by 15 % SDS gel. M-Protein marker; Lanes (1 & 2) - Elute peak fractions.

### 3.1.5 *In-vitro* characterization of FliH (57-258) & FliH (73-258) by Size exclusion chromatography (SEC)

Gel filtration or size exclusion chromatography separates molecules/ proteins based on their molecular mass hence for determination of molecular mass of the protein, the SEC of the purified FliH (57-258) as well as FliH (73-258) were performed on Superdex$^{TM}$ 200. The column was equilibrated with 75 mL of 20 mM Tris pH 8.0 and 100 mM NaCl. 2 mg/mL of the protein was injected onto the column using a 500 µL injection loop. The flow rate used for sample injection and elution was 0.5 mL/min. The protein was injected to the column. Based on elution volume, molecular mass of the protein was calculated.

FliH (57-258) eluted at 14.3 mL while FliH (73-258) eluted at 14.2 mL on Size excusion column. The molecular weight of FliH (57-258) from its sequence is 23 kDa while the predicted molecular weight (based on standard calibration curve in Figure 3.10) was estimated to be 60 kDa. Hence it can be stated that this protein forms at least a dimer in solution. Similarly, the molecular weight of FliH (73-258) from its sequence is 21kDa while the predicted molecular weight from SEC was estimated to be 60 kDa. Therefore, based on the higher predicted molecular weight as well as broad peak size, it can be stated that the FliH most likely exists as an elongated dimer in solution. Since Size exclusion chromatography gives approximate measurement of molecular weight, other sophisticated and sensitive methods like MALS (Multiangle Light Scattering) can be used for further determination of exact molecular weight.

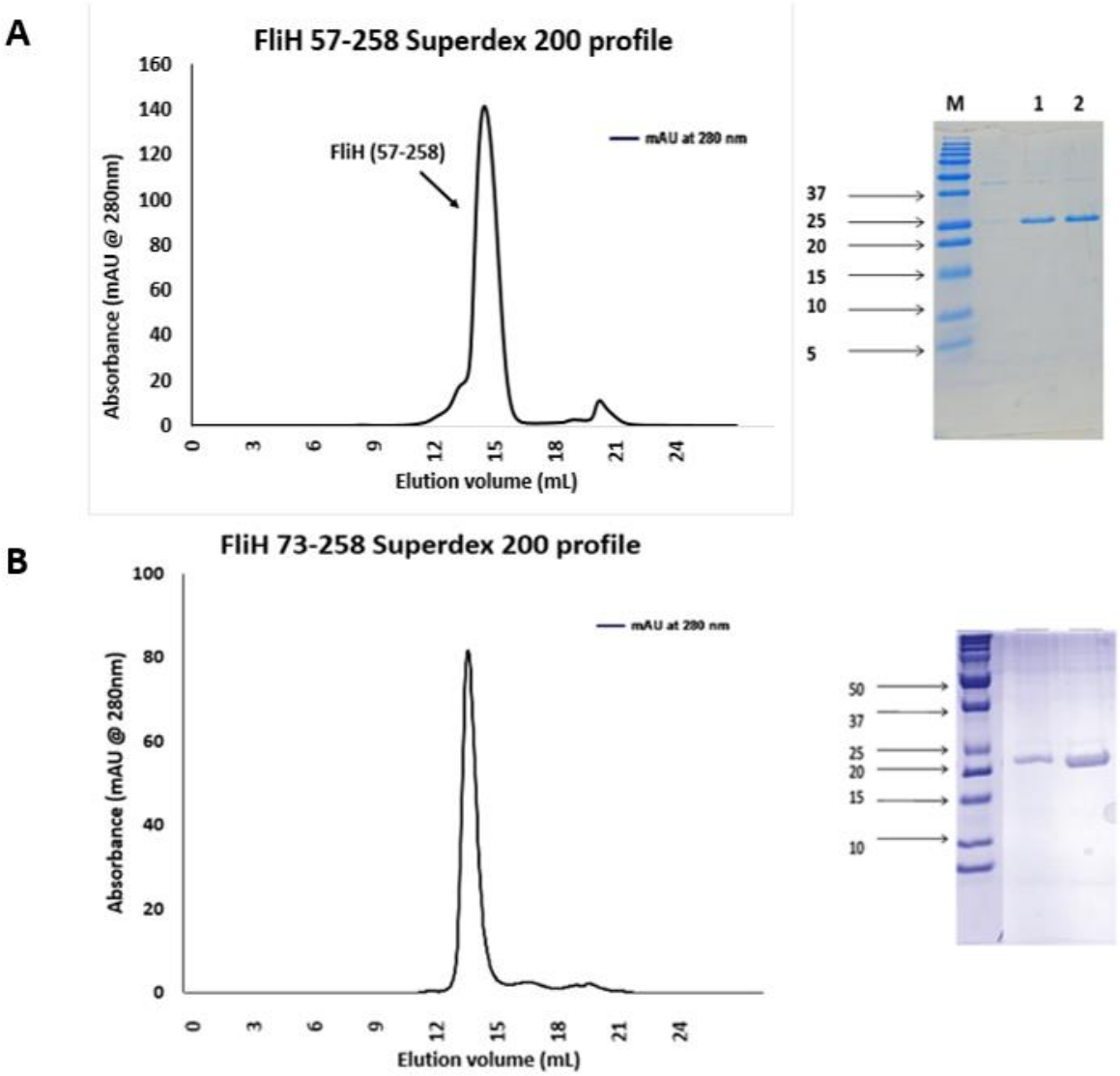


Figure 3.9. **Purification of FliH constructs using a Superdex 200 Size Exclusion column.**

The column was equilibrated with 20 mM Tris pH 8.0 and 100 mM NaCl. 2 mg/mL of the protein was injected onto the column. The flow rate of 0.5 mL/min was used for injection as well as as elution. A) A typical Superdex 200 elution profile for FliH(57-258). FliH(57-258) eluted as a single peak with an elution volume of 14.3 mL (predicted molecular weight 60 kDa). The eluted fractions were analysed by 15% PAGE gel. B) A typical Superdex 200 elution profile for FliH(73-258). FliH(73-258) eluted as a single peak with an elution volume of 14.2 mL (predicted molecular weight 60 kDa). The eluted fractions were analysed by 15% PAGE

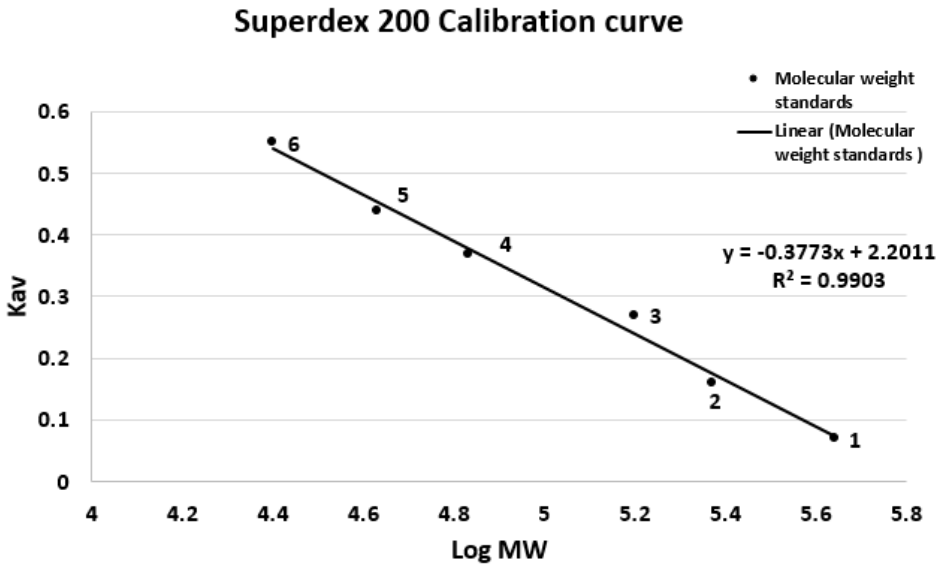


Figure 3.10. **Superdex 200 SEC calibration curve.**

The calibration curve for Superdex 200 SEC.

$K_{av}=V_e-V_0/V_t-V_0$

Where, $V_0$ = void volume, $V_e$= elution volume and $V_t$= column volume; the void volume is the elution volume of Blue Dextran. 1=Ferritin, 2= Catalase, 3= Aldolase, 5= Ovalbumin, 6= Chymotrypsinogen.

Table 3.1. **Calibration data used for Superdex 200 SEC column.**

| | **$V_e$** | **$V_0$** | **$V_t$** | **$K_{av}$** | **MW** | **logMW** |
|---|---|---|---|---|---|---|
| **1. Ferritin** | 9.12 | 8.04 | 24 | 0.07 | 440,000 | 5.64 |
| **2. Catalase** | 10.64 | 8.04 | 24 | 0.16 | 232,000 | 5.37 |
| **3. Aldolase** | 12.40 | 8.04 | 24 | 0.27 | 158,000 | 5.2 |
| **4. Albumin** | 14.00 | 8.04 | 24 | 0.37 | 67,000 | 4.83 |
| **5. Ovalbumin** | 15.12 | 8.04 | 24 | 0.44 | 43,000 | 4.63 |
| **6. Chymotrypsinogen** | 16.80 | 8.04 | 24 | 0.55 | 25,000 | 4.40 |

### 3.1.6 Crystallization of FliH (57-258) & FliH (73-258)

**Crystallization of FliH (57-258)**

Crystallization screening of the FliH 57-258 was performed at 20°C by the hanging-drop vapor-diffusion technique as discussed in Materials and Methods section 2.7. For crystallization following screening kits were used: Wizard I and II (Rigaku reagents), Crystal Screen 1 and Crystal Screen 2 (Hampton Research), JCSG + suite (Qiagen), PACT suite (Qiagen). The screens were set up using the Gryphon Robot (Art Robins Instruments) in 96 well crystal plates.

The drops with crystals were repeated manually using 24 well Linbro plates. Each drop was prepared by mixing 1 μL of the protein solution (15.3 mg/mL) with 1 μL of the reservoir solution and was equilibrated against 1 mL of reservoir solution. Many conditions showed likelihood of growing crystals (Figure 3.10). Very small crystals were observed in drops containing 20% PEG 8000 + phosphate-citrate buffer pH 4.2+ NaCl. We tried to optimize the conditions by varying the precipitant concentration, pH and additives in the original condition of 2.0 M Ammonium sulfate + 0.1M Sodium citrate pH 5.5. Depending on the results of all kits, some more grid screens were set up. Within 4-5 days, small crystals were observed in drops containing 2.0 M Ammonium sulfate + 0.1M Sodium citrate pH 5.5 (Figure 3.11).

We collected the data of the very first crystal (from Ammonium sulfate + Sodium citrate condition) at the 081D-1 small gap undulator beamline, Canadian Light Source, University of Saskatchewan. The crystal diffracted to approximately 3.9 Å. The diffraction pattern of the crystals is shown in Figure 3.11. Images were collected at a crystal to film distance of 290.0 mm. The exposure time per frame was 4 seconds. Each frame was one-degree oscillation. The wavelength was 0.97934 Å. Even after extensive efforts it was impossible to reproduce the crystals of this protein.

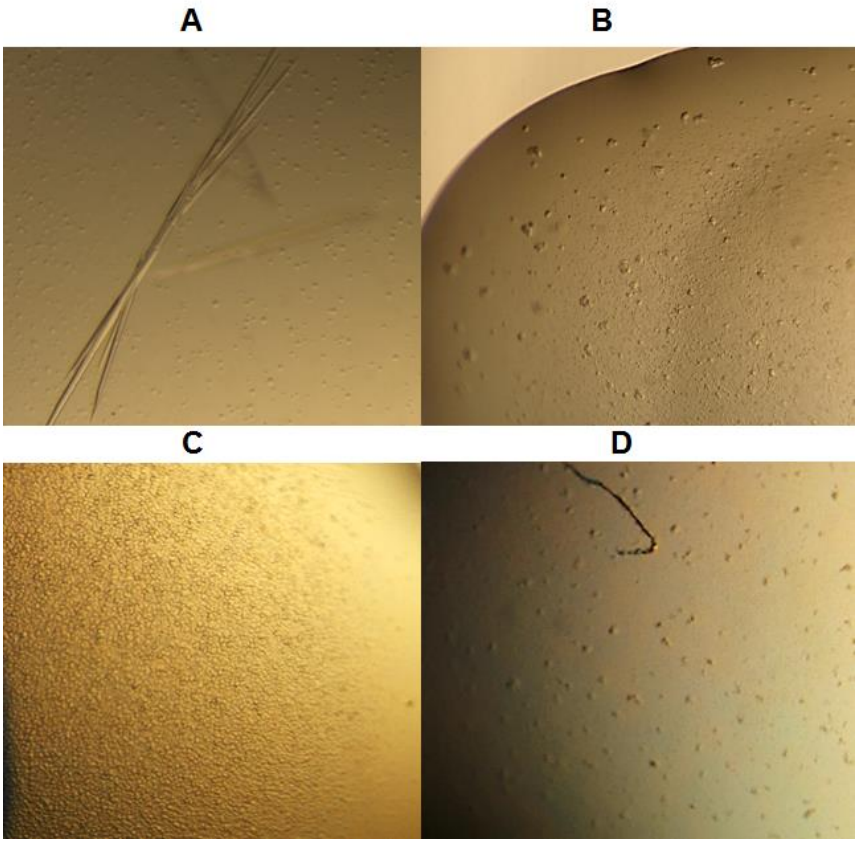


Figure 3.11. **Crystallization of FliH(57-258).**

The crystallization of FliH(57-258) was performed by Hanging Drop Vapor Diffusion technique. The protein concentration used to set up plates was 15.3 mg/mL. The plates were incubated at 20°C in undisturbed condition. Crystals obtained from drops prepared by mixing 1 µL of protein solution (15.3 mg/mL) and 1 µL of reservoir solution containing A) 2.0 M Ammonium sulfate + 0.1M Sodium citrate pH 5.5, B) 20% P8K + 0.1 M CAPS pH 10.5 + 0.2M NaCl, C) 16% P8K + 0.1 M CHES pH 10 + 0.2M NaCl, D) 0.1M BICINE + 10% P6K.

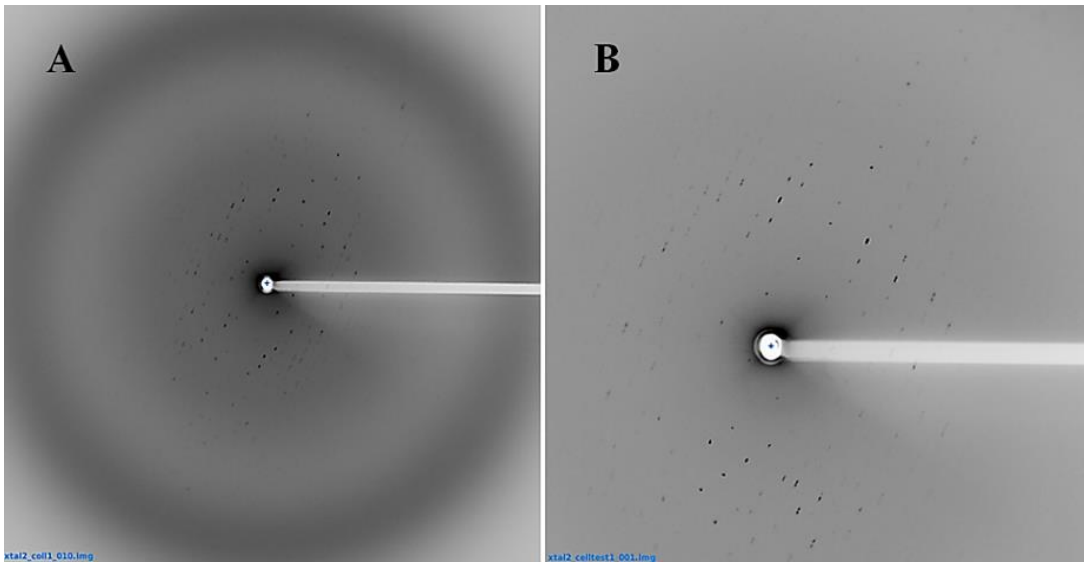


Figure 3.12. **Diffraction images of the FliH (57-258) crystals.**

The diffraction images showing reflections up to 3.9 Å resolution limit. The data was collected at 08ID-1 small gap undulator beamline.

**Crystallization of FliH (73-258)**

Initial crystallization screening for the FliH 73-258 was performed similar to FliH (57-258). Within 1 day small crystals were observed in drops containing 0.1 M Tris pH 8.0 + 20% PEG 6000 (Polyethylene Glycol). This condition was optimized to get better quality crystals (Figure 3.13). Each drop was prepared by mixing 1 µL of the protein solution (10 mg/mL) with 1 µL of the reservoir. We collected the data of these crystals at the 081D-1 small gap undulator beamline, Canadian Light Source, University of Saskatchewan. The maximum diffraction obtained to 8 Å resolution limit. The diffraction pattern of the crystals is shown is Figure 3.14. Images were collected at a crystal to film distance of 140.0 mm. The exposure time per frame was 10 seconds.

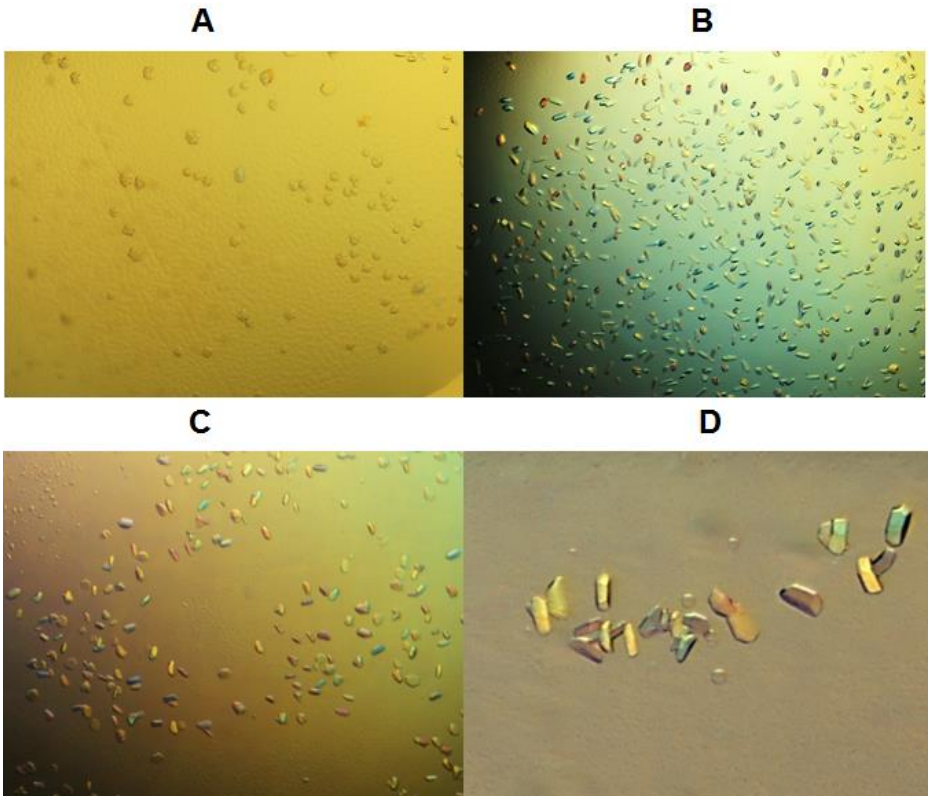


Figure 3.13. **Crystallization of FliH(73-258).**

The crystallization of FliH(57-258) was performed by Hanging Drop Vapor Diffusion technique. The protein concentration used to set up plates was 10 mg/mL. The plates were incubated at 20°C in undisturbed condition. Crystals obtained from drops prepared by mixing 1 µL of protein solution and 1 µL of reservoir solution containing A) 0.1M Tris pH 8.0 + 20% P6K, B) 0.1M Tris pH 7.5 + 21% P6K, C) 0.1M Tris pH 7.9 + 18% P6K, D) 0.1M Tris pH 7.6 + 17% P6K.

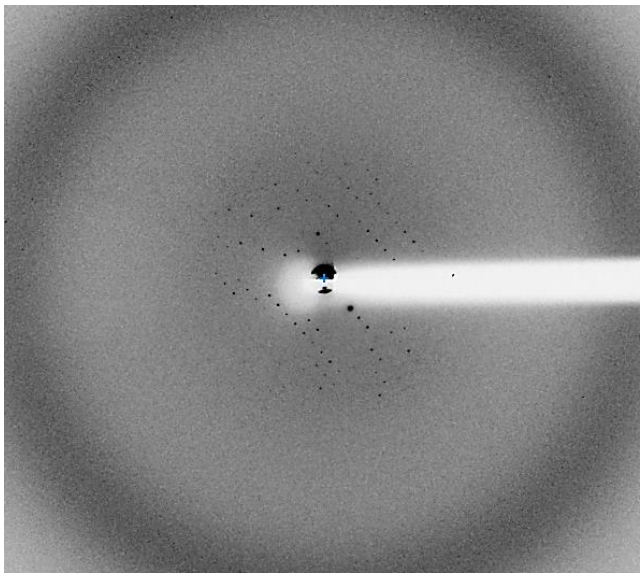

Figure 3.14. **Diffraction image of the crystal of FliH 73-258.**

The diffraction image for the crystals of FliH 73-258 showing reflections up to 8 Å resolution limit. The crystals belong to monoclinic space group, with cell parameters a= 94 Å, b= 157.09 Å, c= 204.00 Å. Data was collected at 08ID-1 small gap undulator beamline.

### 3.2 Characterization of *H. pylori* FliI

#### 3.2.1 Cloning of FliI E193Q mutant

The gene encoding FliI E193Q mutant was cloned in the pGEX-6P-3 vector. Cloning of FliI E/Q mutant was done as described in Materials and Methods section 2.4.10. FliI E193Q mutant gene is 1302 bp or 434 amino acids long. The resultant image of agarose gel after successful cloning and plasmid digestion is shown in Figure 3.15.

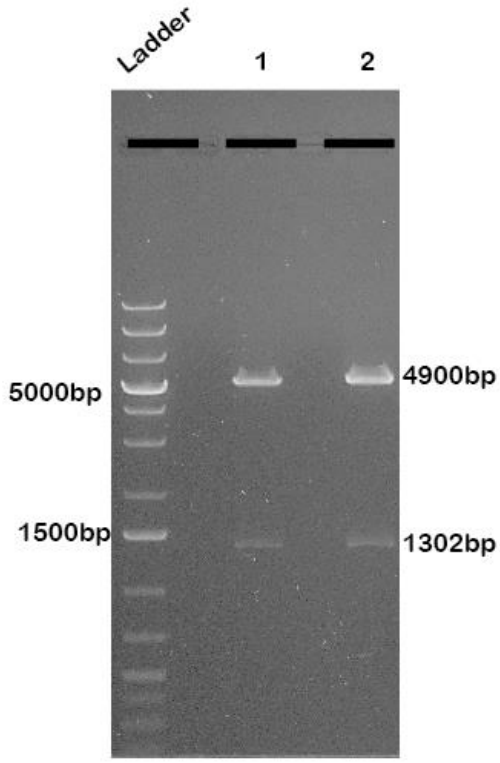


Figure 3.15. **Plasmid digestion of the wild type and mutant FliI.**

Agarose Gel Electrophoresis of the digested wild type FliI Lane (1) and FliIE193Q lane (2) plasmids on 1% Agarose gel with 3µL of HydraGreen dye. 8µL of each the plasmid was mixed with 2 µL of 6X loading dye (Fermentas) and run the gel for 45-50 minutes @100V. 1kb DNA ladder was used as marker. Lane 2 shows the digested FliI E193Q mutant at 1302 bp and pGEX-6P-3 plasmid at 4900 bp.

### 3.2.2 Protein expression and purification of FliI E193Q mutant

FliI E193Q mutant (47 kDa) was purified by affinity purification method as follows.

For expression purposes, *E. coli* BL21 cells were transformed with the FliI E193Q mutant plasmid, and the transformed cells were streaked for single colonies on LB Agar (Amp) plates. The agar plates were incubated at 37°C overnight. A single colony was used to inoculate 50 ml of Hyperbroth™ (Athena Enzyme Systems™), 2.5 mL Glucose Nutrient Mix (Athena Enzyme Systems™), 50 µL 100 ng/ µL of Amp and was incubated overnight at 37°C in a shaking incubator. The following day 2L of Amp Hyperbroth was inoculated with 20 mL of the overnight culture and grown at 37°C shaking incubator until 0.6 AU at $OD_{600}$ nm. For induction of protein expression 0.1 mM IPTG was added to the culture and then 2L of culture was incubated at 20°C overnight. After incubation, the cells were harvested by centrifugation and lysed by cell disruptor, and separated into soluble and insoluble fractions by centrifugation. The pellet was then resuspended in lysis buffer and kept at -20°C for overnight. Samples of the lysate, lysate supernatant and lysate pellet were kept for analysis and run on an SDS-PAGE gel (Figure 3.16). This gel shows that FliI expression has occurred but most of the protein is insoluble and resides in the lysate pellet. The next day the lysate supernatant was subjected to 1$^{st}$ step of GS-affinity chromatography (1$^{st}$ GST) by loading the supernatant on GS-column (25 mL bed volume). The bound GST-FliI fusion protein was easily removed using reduced glutathione in elution buffer. SDS gel image of the samples collected at each step of purification is shown in Figure 3.16. Following buffer exchange and treatment with PSP, the protein sample was again injected onto GS column (2$^{nd}$ GST) to remove free GST. The FliIE193Q mutant protein flows through the column while GST remains bound to the column. The bound GST is then removed from the column using elution buffer. A small amount of sample was analyzed by SDS PAGE (Figure 3.16)

FliI E193Q has a predicted isoelectric point of 6.35, therefore it was decided to further purify this protein by anion exchange chromatography using a SQ column (10mL bed volume). The column was washed with 50 mL of buffer A (50 mM Tris pH 8.0, 1 M NaCl). The cleaved FliIE193Q mutant protein was injected on SQ column at 1 mL/min flow rate and then flow rate was increased to 2 mL/min to wash out unbound protein for another 30 mL (Figure 3.17). A linear gradient of 100-150 mM NaCl over 50 mL of elution buffer was applied to remove any

bound protein. 50 mM Tris pH 8.0, 1 M NaCl was used as the elution buffer for anion exchange chromatography.

Figure 3.17 shows the typical SQ profile for FliI E193Q mutant, which shows that this protein does not bind to the Source-Q sepharose, instead it passes through the column in the void volume. The SDS-PAGE gel in Figure 3.17 indicates that anion exchange chromatography has resulted in purification of protein almost to homogeneity. But two major facts were observed during purification of FliI E/Q mutant. First, the yield of this protein is very low and second, the FliI protein is not very soluble in solution, since precipitation was observed at every stage of purification. The protein continues to precipitate until it reaches 2 mg/mL. Precipitation of *S. typhimurium* FliI has also been observed (Dreyfus *et al*., 1993).

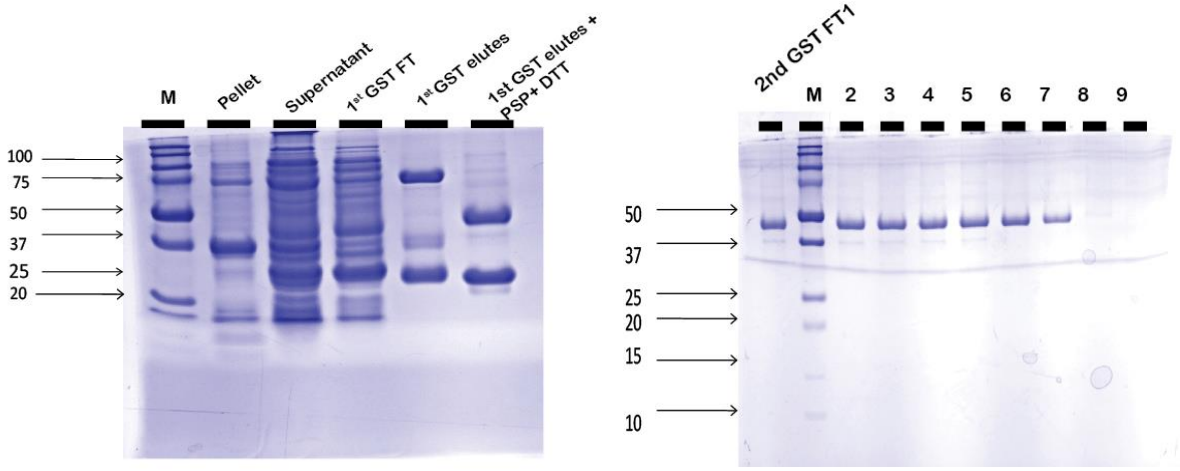


Figure 3.16. **Purification of FliI E193Q mutant.**

Purity of the FliI E193Q at various stages of purification was analyzed by 15% SDS-PAGE. 5 µL of each sample was boiled with 5 µL of loading buffer and loaded on 15% SDS gel. M - Marker, Pellet - lysate pellet, Supernatant - *E. coli* lysate supernatant, 1st GST FT - flow-through fraction obtained after applying lysate supernatant onto the glutathione sepharose column, 1st GST elutes - eluted fractions obtained after applying lysate supernatant onto the glutathione sepharose column, 1st GST elutes+ PSP+ DTT, 2nd GST FT (1-9) – Flow through fractions obtained from the 2nd GS purification step showing the presence of purified FliI E193Q mutant protein (47kDa).

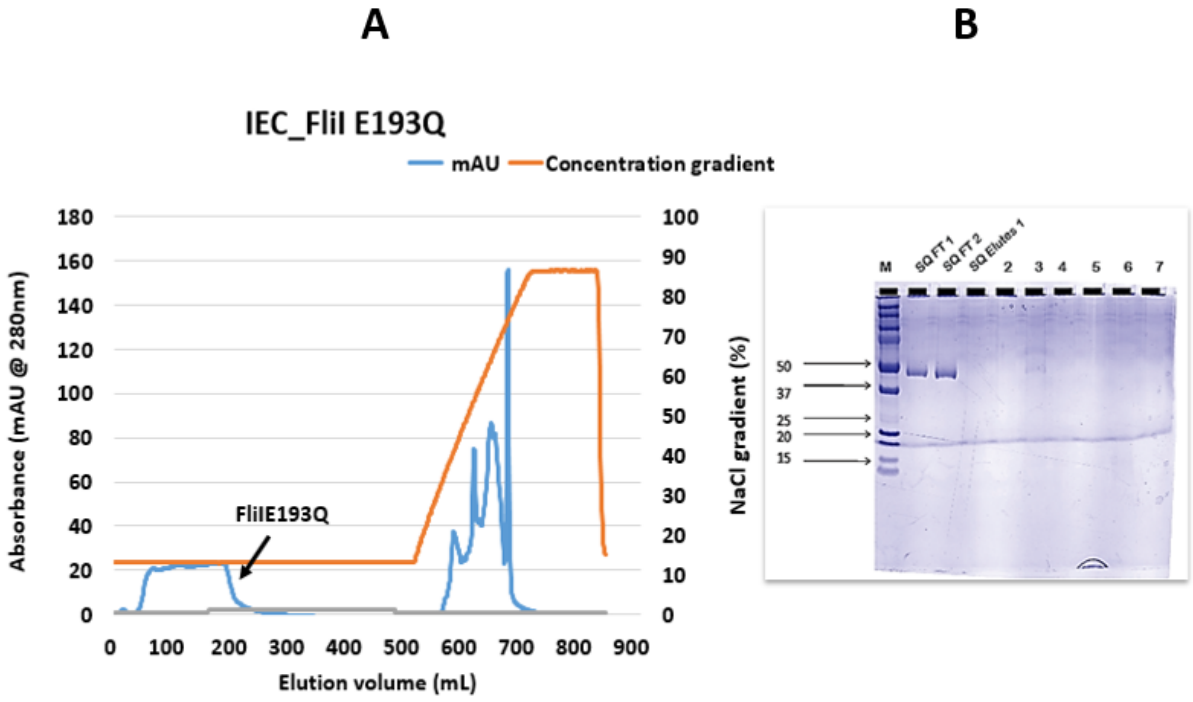


Figure 3.17. **Purification of FliI E193Q using Source-Q Anion Exchange Chromatography.**

A) Elution profile for purification of FliI E193Q. The column was equilibrated in 50 mL of buffer A (50 mM Tris pH 8.0, 1 M NaCl); the flow rate of injection was 1 mL/min. The bed volume of the column was 10 mL. Protein elution was monitored by absorbance (mAu) at 280 nm. B) SDS-PAGE of FliI E193Q anion exchange chromatography. Flow through and elute peak fractions were analyzed by 15 % SDS gel. M-Protein marker; SQ FT- Ion exchange flow through fractions, SQ elutes (1-8) - Elute fractions.

### 3.2.3 *In-vitro* characterization of *H. pylori* FliI E193Q mutant by Size exclusion chromatography

SEC of the purified FliI mutant was performed on a Superdex$^{TM}$ 200 column. The column was equilibrated with 50 mM $Na_3PO_4$ pH 8.0, 150 mM NaCl and 5 mM $MgCl_2$. The purified protein was incubated in buffer containing 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 5 mM $MgCl_2$, 100 mM Urea, 10 mM AMP-PNP and 1 mM DTT for overnight at 4°C on a rocker. The flow rate used for sample injection and elution was 0.5 mL/min. Based on elution volume, molecular mass of the protein was calculated. *Salmonella* FliI hexamerization has been observed in the presence of non-hydrolysable ATP analog by electron microscopy (Kazetani *et al*., 2009), hence it was decided to use AMP-PNP. Based on gel filtration results for *H. pylori* FliI mutant (Figure 3.18), it was predicted to be a monomer in solution.

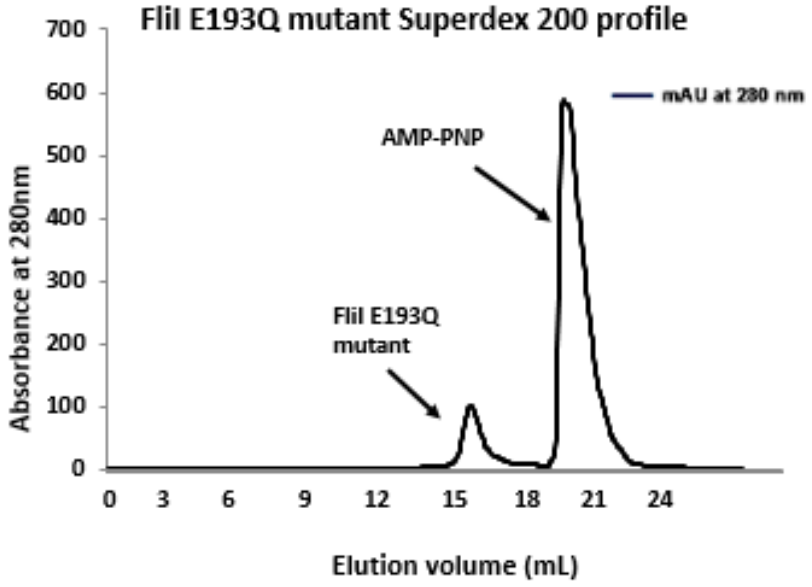


Figure 3.18. **Purification of FliI E193Q mutant using a Superdex 200 Size Exclusion column.**

The column was equilibrated with 50 mM Sodium phosphate pH 8.0, 150 mM NaCl and 5 mM $MgCl_2$. 2 mg/mL of the protein was injected onto the column. The flow rate of 0.5 mL/min was used for injection as well as for elution. A typical Superdex 200 elution profile for FliI E193Q mutant. FliI E193Q mutant eluted as a single peak with an elution volume of 15.36 mL. AMP-PNP eluted at 20 mL(predicted molecular weight 48 kDa).

The predicted molecular weight of FliI E/Q mutant from gel filtration chromatography (Standard curve in Figure 3.10) was found to be 48 kDa. Despite conditions presumably favorable for hexamerization, FliI E193Q mutant did not form hexamers. There could be three reasons for lack of hexamerization. First, the experimental conditions are not sufficient enough for FliI mutant hexamerization. FliI could be forming hexamer in very small concentration and size exclusion chromatography is not sensitive enough to detect the very small amount of hexamer. Second reason could be that the protein is not folded properly. The third possible reason is that FliI requires another protein/chaperone for its hexamerization. Recently it has been observed that FliI requires another general chaperone called FliJ for its hexamerization. FliJ promotes FliI hexamerization by binding at the center of FliI hexametric ring structure (Ibuki *et al*, 2013). Further experimental evidence is required in order to confirm the oligomerization state of *H. pylori* FliI E193Q mutant.

### 3.2.4 Crystallization trails for FliI (E/Q) mutant

Crystallization trials for FliI E/Q mutant was carried out in presence of AMP-PNP in similar way as FliH crystallization, but no crystals were obtained for this protein.

## 3.3 Making & characterization of FliH:FliI (E/Q) mutant complex

We decided to check the interaction of FliH and FliI, since both the proteins could be over-expressed and purified. Individual proteins FliH and FliI were purified by FPLC as mentioned earlier. Both the proteins were mixed in a 1:1 molar ratio of the molecules. The protein mixture was added to 5 mM $MgCl_2$, 1 mM AMP-PNP and 2 mM DTT (final concentrations). Final volume was made up to 500 uL using 1 x PBS. The mixture was incubated for 30 min at room temperature on the rocker and injected on the Superdex 200 with the flow rate of 0.5 ml/min. The Gel filtration profile of the FliH:FliI complex is shown in Figure 3.19A. The fractions were collected and concentrated. The samples were analyzed by loading on a15% SDS gel Figure 3.19B.

Using the Superdex 200 calibration curve (Figure 3.10), the estimated molecular weight of peak 1 fractions was found to be 75 kDa. Since molecular weight of monomeric FliI E/Q mutant is 47 kDa and that of FliH 57-258 is 21 kDa, there is a possibility of the formation of $FliH_2$:FliI complex. The solubility issues of the FliI mutant made the FliH:FliI complex difficult to concentrate and purify. The protein solution was precipitating while concentrating possibly

due to poor solubility of the FliI. It is difficult to determine the exact molecular weight of the complex using SEC, hence more sensitive techniques like MALS will be used in the future.

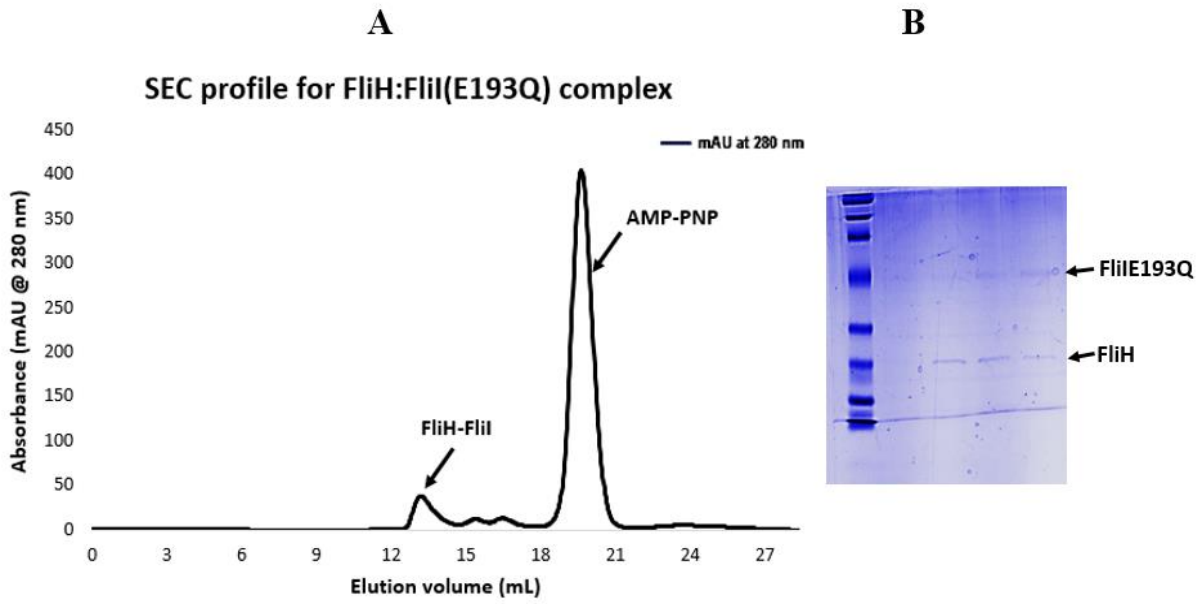


Figure 3.19. **SEC profile of the FliH:FliI E/Q mutant complex.**

**A**. Superdex 200 elution profile of mixture of FliH 57-258 and FliI E/Q mutant complex. Both the proteins were at concentration of 2 mg/mL before they were mixed. The mixture contained 500 µg of FliH 57-258 and 370 µg of FliI E/Q mutant, such that the proteins were at 1:1 molar ratio. The protein mixture was incubated with 5 mM $MgCl_2$, 1 mM AMP-PNP and 2 mM DTT at room temperature on rocker for 30 min, followed by injection onto Superdex 200 column. **B**. 15% SDS-PAGE of the fractions from 1st eluted peak. Eluted fraction size- 2 mL/tube. 5 µL of each fraction was mixed with 5 µL of loading buffer and the sample was loaded on the gel (Predicted molecular weight 75 kDa).

### 3.4 Characterization of *H. pylori* HP0256

#### 3.4.1 Cloning of HP0256

The gene encoding HP0256 was cloned in the pGEX-6P-3 vector. Cloning of HP0256 was done as described in Materials and Methods section 2.2. HP0256 gene is 441 bp or 147 amino acids long. The restriction digested plasmid was run on agarose gel (Figure 3.20).

#### 3.4.2 Protein expression and purification of HP0256

Purification of HP0256 (16 kDa) was carried out by Affinity Chromatography using the same protocol as for other proteins. The SDS gel image of the purified protein is shown in figure 3.21. The purified protein was concentrated to 10 mg/mL and used for SEC by Superdex 200 column. The column was equilibrated with 20 mM Tris pH 8.0 and 100 mM NaCl. The flow rate used for sample injection and elution was 0.5 mL/min. Based on elution volume (30.8 mL), the predicted molecular mass of the protein was calculated (Figure 3.22). The molecular weight of FliJ from its sequence is 16 kDa while the molecular weight of this protein from gel filtration was estimated to be 24 kDa hence it can be presumed that FliJ remains as a monomer in solution.

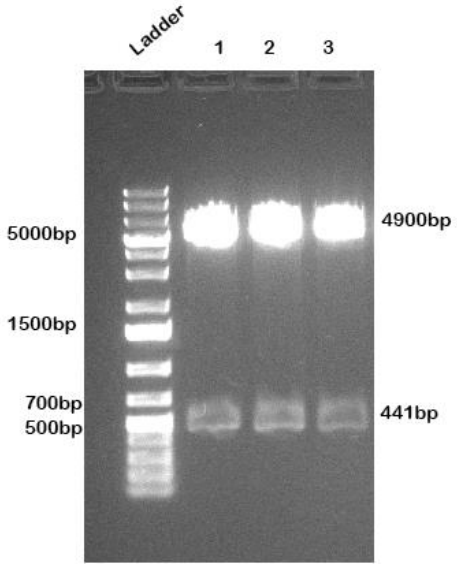


Figure 3.20. **Plasmid digestion of the HP0256 gene.**

Agarose Gel Electrohporesis of the HP0256 plasmids (digested with *Bam*HI and *Eco*RI, 2 µL @20,000 U/mL each) on 1% Agarose gel with 3µL of HydraGreen dye. 8µL of each plasmid was mixed with 2 µL of 6X loading dye (Fermentas) and run the gel for 45-50 minutes @100V. 1kb DNA ladder was used as marker. Lanes 1-3 shows the digested HP0256 at 441 bp and pGex-6P-3 plasmid at 4900 bp.

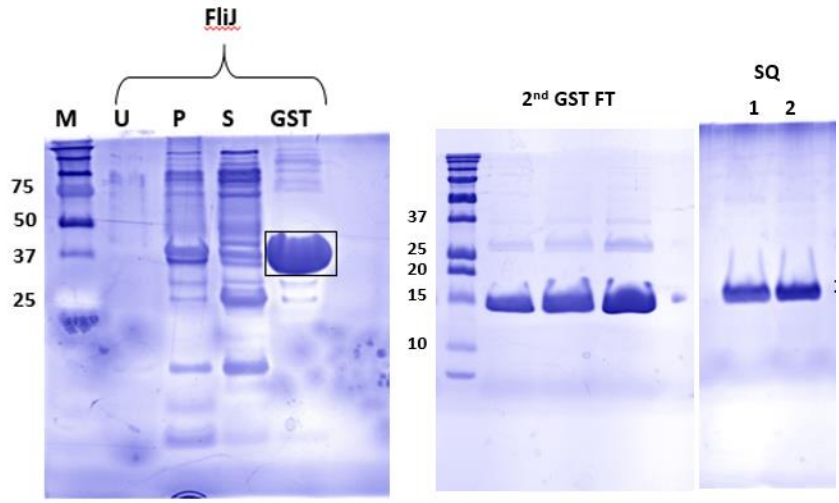


Figure 3.21. **Purification of HP0256 using Source-Q Anion Exchange Chromatography.**

Purity of HP0256 at various stages of purification was analyzed by 15 % SDS-PAGE. M-Marker; U – uninduced sample; P – *E. coli* lysate pellete; S - *E. coli* lysate supernatant; GST - The eluted GST-HP0256 fusion protein (42 kDa) fractions obtained after applying lysate supernatant onto the glutathione sepharose column; 2nd GST FT - flow-through fractions from the 2nd GST purification HP0256 (16 kDa); SQ – Elutes obtained after applying cleaved HP0256 to the ion-exchange column.

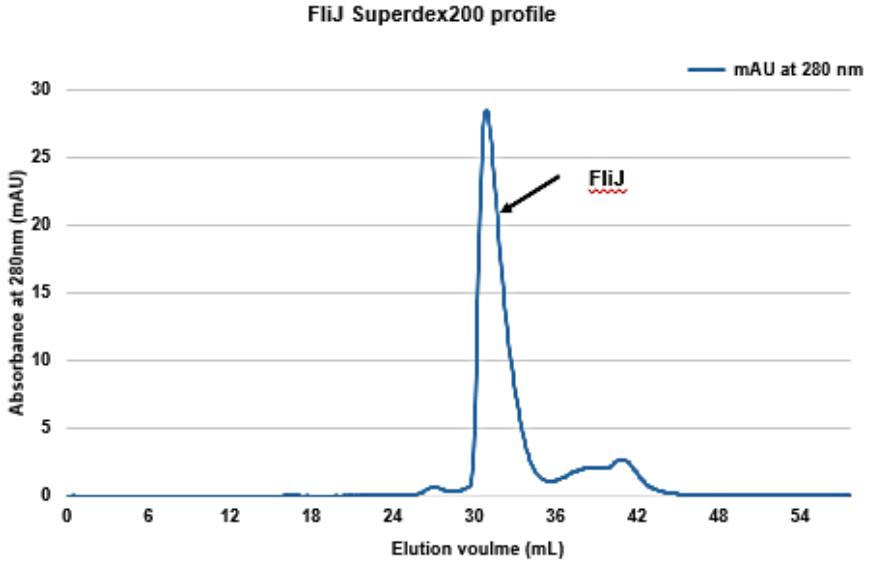


**Figure 3.22. Purification of HP0256 using a Superdex 200 Size Exclusion column.**

The cloumn was equilibrated with 20 mM Tris pH 8.0 and 100 mM NaCl. 2 mg/mL of the protein was injected onto the column. The flow rate of 0.5 mL/min was used for injection as well as for elution. A typical Superdex 200 elution profile for HP0256 is shown. HP0256 eluted as a single peak with an elution volume of 30.8 mL (predicted molecular weight 24 kDa).

### 3.4.3 Crystallization trials for HP0256

The initial screen for crystallization was carried out using a Gryphon robot. Crystallization screening of the HP0256 was performed at 20°C by the hanging-drop vapor-diffusion technique by using the following screening kits: Wizard I and II (Rigaku reagents), Crystal Screen 1 and Crystal Screen 2 (Hampton Research), JCSG + suite (Qiagen), PACT suite (Qiagen). Within 1 day small crystalline precipitates were observed in a few drops (Figure 3.23). Optimization of crystallization is currently in progress.

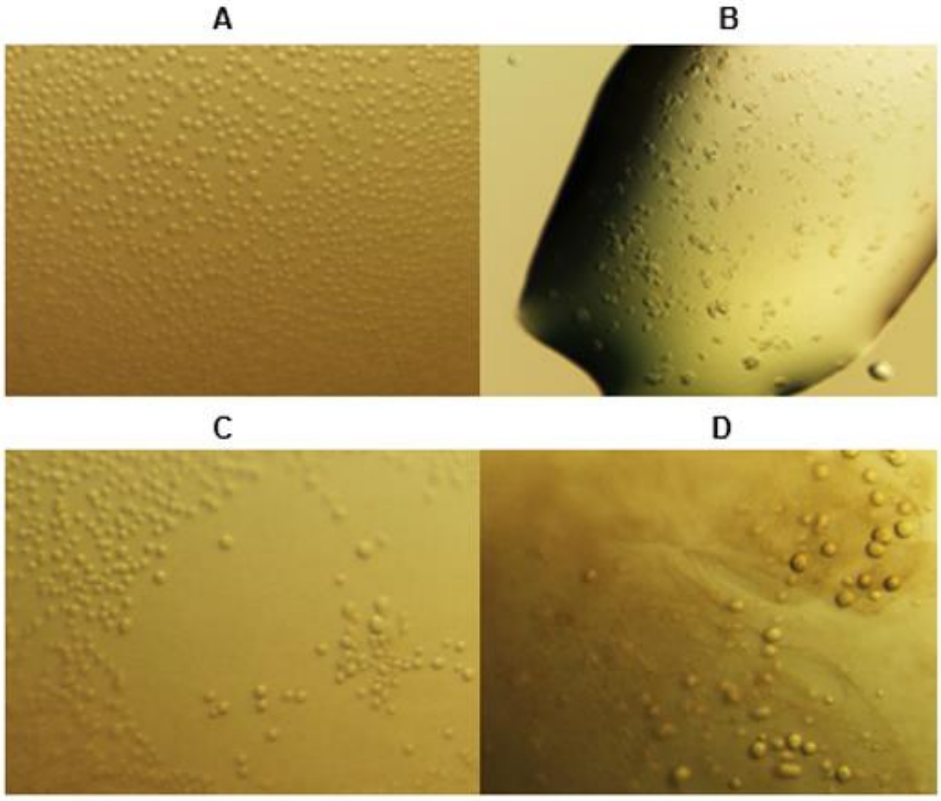


Figure 3.23. **Crystallization of HP0256.**

The crystallization of HP0256 was performed by Hanging Drop Vapor Diffusion technique. The protein concentration used to set up plates was 8.5 mg/mL. The plates were incubated at 20°C. Crystalline precipitates/phase seperation obtained from drops prepared by mixing 1 µL of protein solution and 1 µL of reservoir solution containing A) 10% w/v PEG 8000 + 0.1 M Tris pH 7.0 + 0.2 M $MgCl_2$, B) 0.1 M MES pH 6.0 + 0.2 M $Zn(OAc)_2$ + 12% Isopropanol, C) 10% w/v PEG 8000 + 0.1 M CHES pH 9.5 + 0.2 M NaCl, D) 30% v/v PEG 400 + 0.1 M CAPS pH 10.5

### 3.5 Characterization of FliS and $FlaA_2$ (463-488)

#### 3.5.1 Cloning of FliS, $FlaA_2$ (463-488)

Cloning of FliS (378bp), $FlaA_2$ (463-488) (138bp) was done from *H. pylori* NCTC 11637genomic DNA using standard cloning protocol. After successful cloning the concentration of plasmids were found to be 188.5 ng/µL and 147.5 ng/µL respectively. The resultant image of agarose gel after successful cloning and plasmid digestion is shown in Figure 3.24.

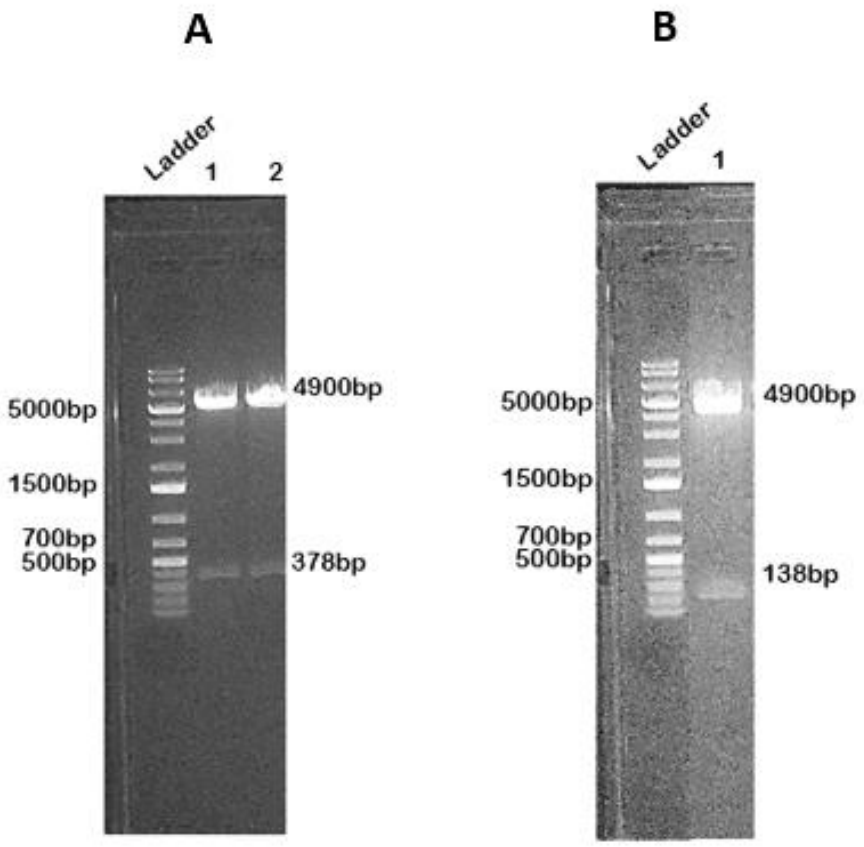


Figure 3.24. **Plasmid digestion of the FliS, FlaA2 (463-488) gene.**

Agarose Gel Electrohporesis of the FliS, FlaA2 (463-488) (digested with *Bam*HI and *Eco*RI, 2 µL @20,000 U/mL each) plasmids on 1% Agarose gel with 3µL of HydraGreen dye. 8µL of each plasmid was mixed with 2 µL of 6X loading dye (Fermentas) and run the gel for 45-50 minutes @100V. 1kb DNA ladder (L) was used as marker. **A)** Lanes 1&2 shows the digested FliS at 378 bp and pGex-6P-3 plasmid at 4900 bp, **B)** Lane 1 shows the digested $FlaA_2$ (463-488) at 138bp and pGex-6P-3 plasmid at 4900 bp.

### 3.5.2 Expression and purification of FliS

FliS was purified using the same method as described in Materials and Methods section 2.4. Figure 3.25 shows the SDS gel of the fractions after successive stages of purification. FliS has molecular weight of 14.5 kDa (126 amino acids) and predicted molecular weight of GST- FliS fusion protein is approximately 40.4 kDa. The fusion protein remains bound to GS column which was then was then eluted using 50 mL elution buffer (50 mM Tris HCl pH 8.0, 10 mM reduced glutathione). Figure 3.25 shows the fractions analyzed by SDS PAGE. The GST tag was cleaved in similar way as explained earlier in this study. The sample was subjected to 2nd round of purification by injecting onto GS column. The FliS comes out as flow-through.

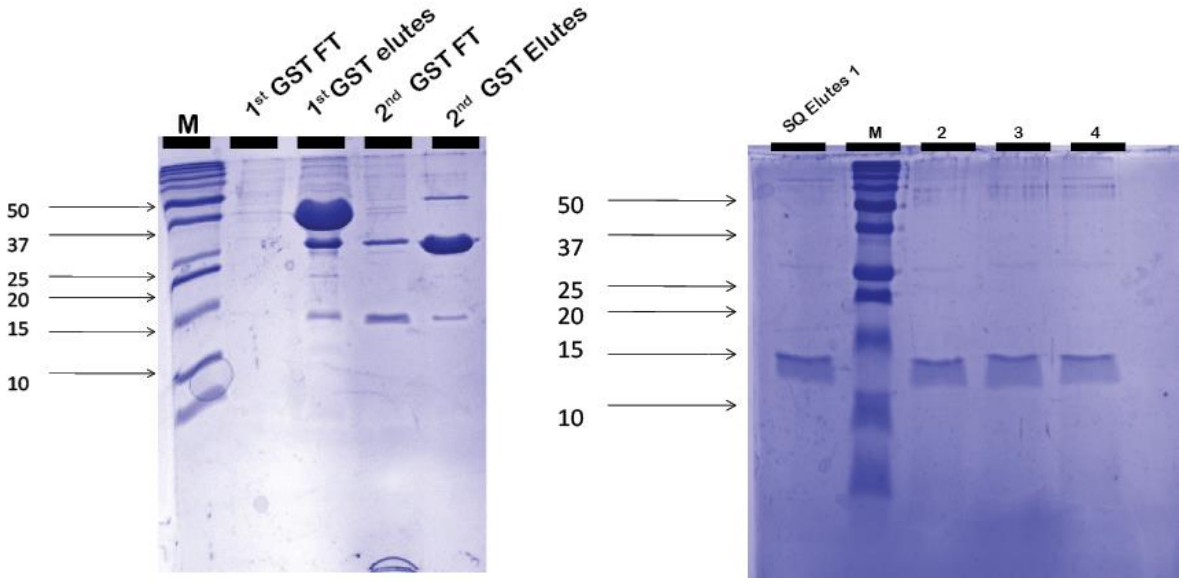


Figure 3.25. **Purification of FliS by GS-Affinity purification method.**

Purity of the FliS at various stages of purification was analyzed by 15% SDS-PAGE. 5 µL of each sample was boiled with 5 µL of loading buffer and loaded on 15% SDS gel. M - Marker, Pellet - lysate pellet, Supernatant - *E. coli* lysate supernatant, 1st GST FT - flow-through fraction obtained after applying lysate supernatant onto the GS column, 1st GST elutes - eluted fractions obtained after applying lysate supernatant onto the GS column, 2nd GST FT – Flow through fractions obtained from the 2nd GS purification step showing the presence of purified FliS protein at 14.5 kDa, 2nd GST elutes- eluted GST fractions at 26 kDa, SQ elutes (1-4) - Eluted FliS fractions after applying FliS to Ion exchange column.

FliS has predicted isoelectric point of 5.03, therefore it was decided to further purify this protein by anion exchange chromatography using SQ column. The column was washed with 50 mL of buffer A (50 mM Tris pH 8.0, 1 M NaCl). The cleaved FliS protein was injected on the SQ column at 1 mL/min flow rate and then the flow rate was increased to 2 mL/min to wash out unbound protein for another 30 mL. The FliS was eluted by applying a linear gradient of 100-150 mM NaCl over three column volumes using buffer A and elution buffer B (50 mM Tris-HCl pH 8.0, 2 M NaCl). The eluted fractions were analyzed by SDS-PAGE gel (Figure not shown).

Appropriate fractions containing FliS were pooled and concentrated to 10 mg/mL using an Amicon® Ultra-15 centrifugal filter device, 3 kDa Molecular Weight Cut Off (MWCO) and utilized for crystallization studies. Purified FliS was analyzed by SDS-PAGE gel (Figure 3.26)

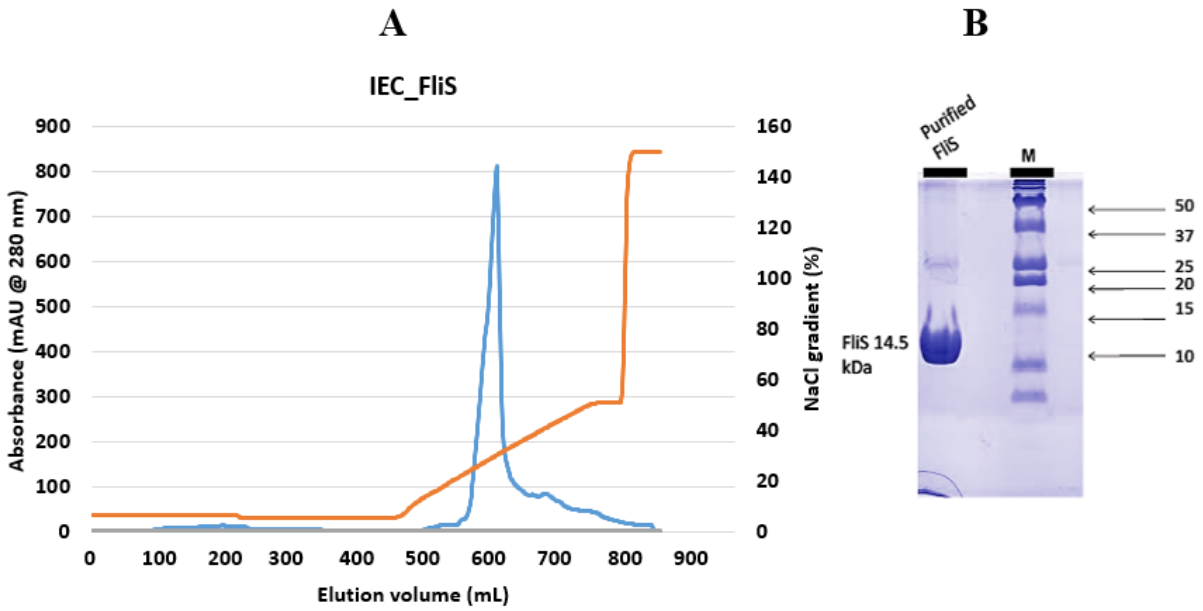


Figure 3.26. **Purification of FliS using Source-Q Anion Exchange Chromatography.**

A) Elution profile for purification of FliS. The column was equilibrated in 50 mL of buffer A (50 mM Tris pH 8.0, 1 M NaCl); the flow rate of injection was 1 mL/min. The bed volume of the column was 10 mL. Protein elution was monitored by absorbance (mAu) at 280 nm. B) The appropriate fractions were pooled and concentrated to 10 mg/mL and were analyzed by 15 % SDS-PAGE gel.

### 3.5.3 *In-vitro* characterization of FliS by Size exclusion chromatography

Gel filtration or size exclusion chromatography separates molecules/ proteins based on their molecular mass hence for determination of molecular mass of the protein, the SEC of the purified FliS was performed on Superdex$^{TM}$ 75. The column was washed with 75 mL of water and then equilibrated with 75 mL of Buffer A 10 mM Tris pH 8.0. 2 mg/mL of the protein was injected onto the column using a 500 µL injection loop. The flow rate used for sample injection and elution was 0.5 mL/min. The FliS eluted at 13.2 mL (Figure 3.27). Based on the elution volume, molecular mass of the protein was calculated. The molecular weight of FliS from its sequence is 14.5 kDa while the predicted molecular weight from standard curve (Figure 3.28) was estimated to be 17 kDa. The *Salmonella* FliS have been shown to form dimer in solution (Auvray *et al*., 2001). Yet, from the SEC results it can be stated that *H. pylori* FliS is present as a monomer in solution.

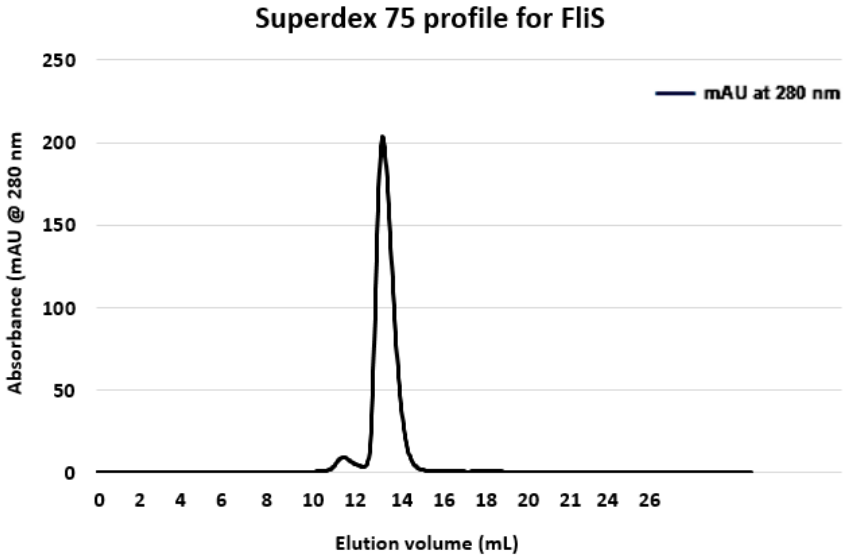


Figure 3.27. **Purification of FliS using a Superdex 75 Size Exclusion column.**

The cloumn was equilibrated with Buffer A 10 mM Tris pH 8.0, 150 mM NaCl. 2 mg/mL of the protein was injected onto the column. The flow rate of 0.5 mL/min was used for injection as well as for elution. A typical Superdex 75 elution profile for FliS. FliS eluted as a single peak with an elution volume of 13.2 mL (predicted molecular weight 17 kDa).

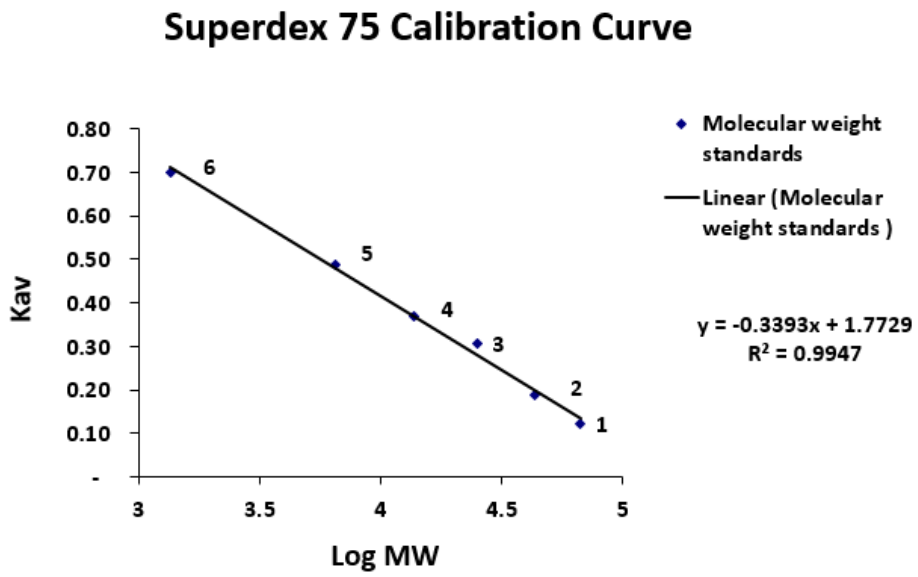


Figure 3.28. **Superdex 75 SEC calibration curve.**

The calibration curve for Superdex 75 SEC.

$K_{av}=V_e-V_0/V_t-V_0$

Where, $V_0$ = void volume, $V_e$= elution volume and $V_t$= column volume; the void volume is the elution volume of Blue Dextran. 1= Albumin, 2= Ovalbumin, 3= Chymotrypsinogen A, 4= Ribonuclease, 5= Aprotinin, 6= Vitamin $B_{12}$.

Table 3.2. **Data used for Calibration of Superdex 75 column**

| | **$V_e$** | **$V_0$** | **$V_t$** | **$K_{av}$** | **MW** | **logMW** |
|---|---|---|---|---|---|---|
| **1. Albumin** | 9.84 | 7.84 | 24 | 0.12 | 67000 | 4.826 |
| **2. Ovalbumin** | 10.88 | 7.84 | 24 | 0.19 | 43000 | 4.633 |
| **3. Chymotrypsinogen A** | 12.8 | 7.84 | 24 | 0.31 | 25000 | 4.397 |
| **4. Ribonuclease** | 13.8 | 7.84 | 24 | 0.37 | 13700 | 4.136 |
| **5. Aprotinin** | 15.73 | 7.84 | 24 | 0.49 | 6500 | 3.812 |
| **6. Vitamin $B_{12}$** | 19.15 | 7.84 | 24 | 0.70 | 1355 | 3.131 |

### 3.5.4 Crystallization of FliS

The purified FliS was concentrated and various concentrations starting from 4 mg/mL to 10 mg/mL were used for crystallization. Plates with 4 mg/mL protein concentration shoed the presence of crystals hence it was decided to carry out crystallization for FliS at 4 mg/mL. The initial screen for crystallization was carried out using a Gryphon robot. Crystallization screening of the FliS was performed at 20°C by the hanging-drop vapor-diffusion technique by using the following screening kits: Wizard I and II (Rigaku reagents), Crystal Screen 1 and Crystal Screen 2 (Hampton Research), JCSG + suite (Qiagen), PACT suite (Qiagen). Within 1 day small showers of crystals were obtained in drop containing 10% v/v Iso-Propanol + 0.1 M MES pH 6.0 + 0.2 M $Ca(OAc)_2$ (Figure 3.29A) and 0.1 M HEPES-Na pH 7.5 + 10% v/v Iso-Propanol + 20% PEG 4000. For getting good crystal packing, slow and steady growth of the crystal is very important. We collected the dataset for a few crystals at Canadian light source, U of S. Some of the data sets were not useful for solving crystals due to poor quality. One of the crystals showed good diffraction up to 2Å resolution but we could not process the diffraction images using HKL 2000 software (Otwinowski and Minor, 1997). In the case of FliS, crystals were growing very fast therefore we decided to use Al's Oil (Hampton research), MPD (2-methyl-2, 4-pentanediol), PEG 400, Glycerol to optimize the crystal growth. Various parameters like protein concentration, pH of the buffer, salt concentration etc. were used to optimize the crystal growth. Al's Oil and 4% MPD as well as HXD (1, 6-Hexanediol) (Sigma-Aldrich) helped to slow down the growth of crystals. Incubation at lower temperature (15°C) helped to get good crystals. Finally, good quality crystals were obtained with drops containing 0.1 M MES + 0.2 M $Ca(OAc)_2$ + 7% (v/v) 2-propanol + 4% MPD using 4.5 mg/mL protein concentration. The drops were prepared by mixing 3 µL of protein solution with the 3 µL of reservoir solution. The reservoir solution was covered with 400 µL Al's oil. Within 4-5 days small crystals appeared and they grew slowly within next couple of days. Crystals from several promising conditions were subjected to diffraction. The diffraction experiments were carried out at the Canadian Light Source (CLS) at the University of Saskatchewan. The crystals yielded reflections up to 2.1 Å resolution (Figure 3.30).

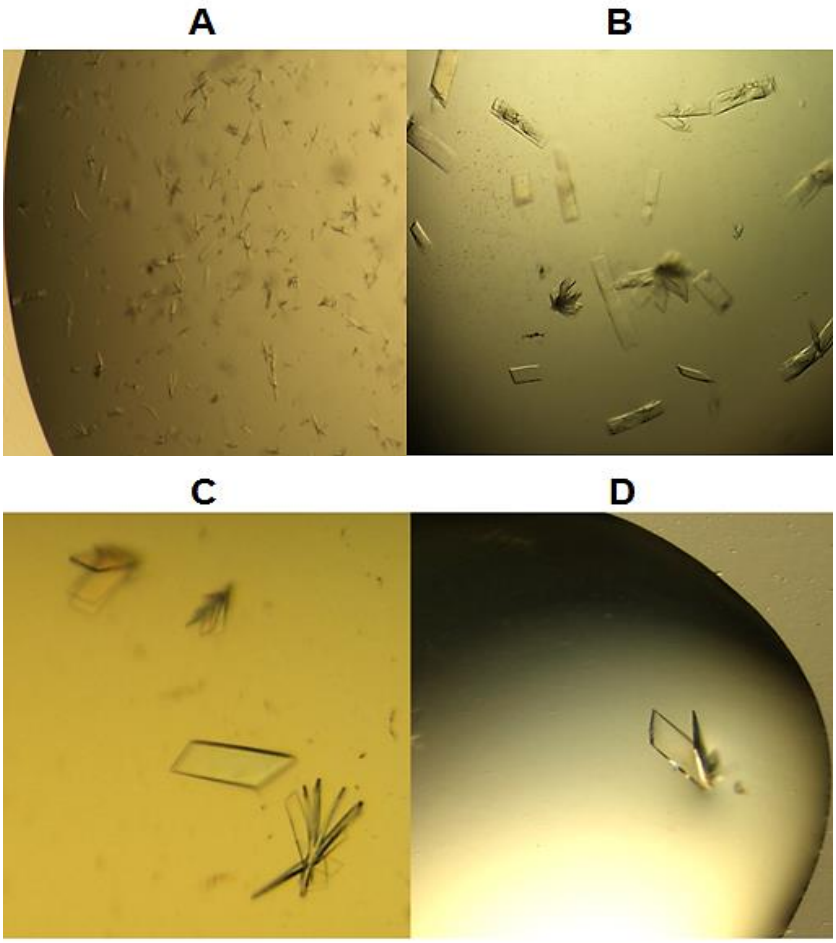


Figure 3.29. **Crystallization of FliS.**

The crystallization of FliS was performed by Hanging Drop Vapor Diffusion technique. The plates were incubated at 15°C. Crystals obtained from drops prepared by mixing 3 µL of protein solution and 3 µL of reservoir solution. A) Protein concentration - 3.8 mg/mL; incubation temperature 20°C reservoir-10% (v/v) 2-propanol + 0.1 M MES pH 6.0 + 0.2 M $Ca(OAc)_2$, B) Protein concentration – 5.8 mg/mL; reservoir - 6% (v/v) 2-propanol + 0.1 M MES pH 6.5 + 0.2 M $Ca(OAc)_2$ + 4% MPD, C) Protein concentration – 4.5 mg/mL; reservoir - 7% (v/v) 2-propanol + 0.1 M MES pH 6.3 + 0.2 M $Ca(OAc)_2$ + 5% MPD; D) Protein concentration – 4.5 mg/mL; reservoir - 7% (v/v) 2-propanol + 0.1 M MES pH 6.3 + 0.2 M $Ca(OAc)_2$ + 5% MPD.

### 3.5.5 Crystal structure of FliS

The diffraction data was processed by Michel Fodje using XDS software (Kabsch, 2010). For refinement and structure solving, Phaser crystallographic software was used (McCoy *et al*., 2007). The structure of FliS (101 residues) has been refined. The data showed reflections up to 2.2 Å resolution, but for solving structure we used reflections only upto 2.45 Å resolution. The structure contains two molecules in the asymmetric unit. The crystal was found to have a space group of P $2_1$ $2_1$ $2_1$ with unit cell parameters a=33.86, b=82.63, c=92.26 and α=β=γ=90°. The *H. pylori* FliS (PDB ID 3IQC) was used as a search model for solving the structure by molecular replacement. Initially a poly-alanine model was used for residues whose density map was not observed. All the residues whose side chain density was not observed, were replaced by an Alanine residue and then the model was refined to see if the density comes back. The residues with missing backbone density were omitted from final structure. The process was repeated until $R_{free}$ and $R_{work}$ values considerably decreased (0.46 and 0.40). For proteins R-factor ranges between 0.6 (for imperfectly refined model) to 0.2 (for well refined model). The final model is shown in Figure 3.32.

Phenix.Xtriage was used to check the quality of the data. It calculates various statistical parameters (like solvent content and Matthews coefficient, twinning and NSC, data strength and completeness etc.) to indicate the quality of the data. In this case the presence of twinning (Figure 3.31) or weak translational pseudosymmetry (from Patterson analysis) can be responsible for slightly high R-value. The data also showed the presence of anisotropy. We could not improve the $R_{free}$ and $R_{work}$ values (beyond 0.46 and 0.40 respectively) after refinement.

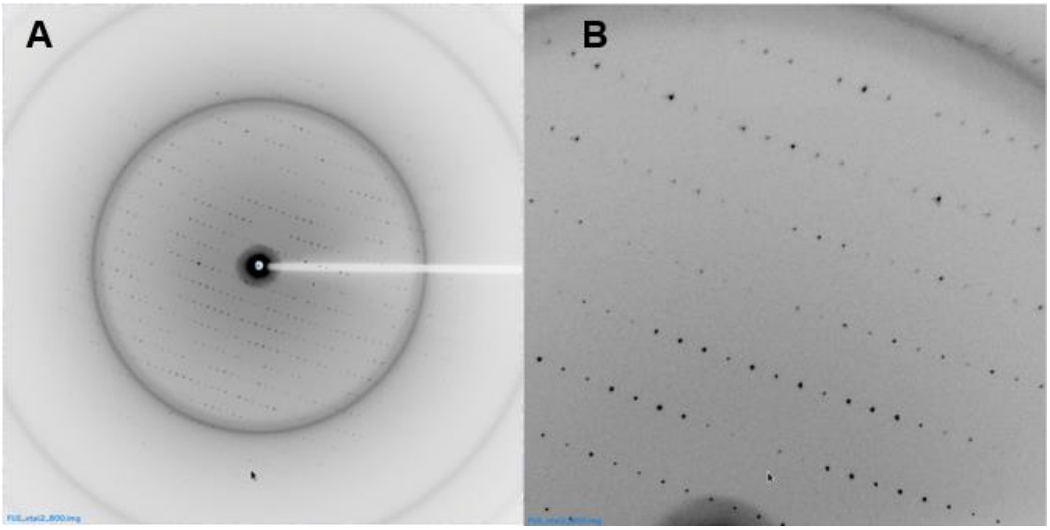


Figure 3.30. **The X-ray diffraction pattern of a FliS crystal.**

**A**. The diffraction image showing reflections up to 2.2Å resolution limit. **B**. Zoomed image of the diffraction pattern. The data was collected at 081D-1 small gap undulator beamline.

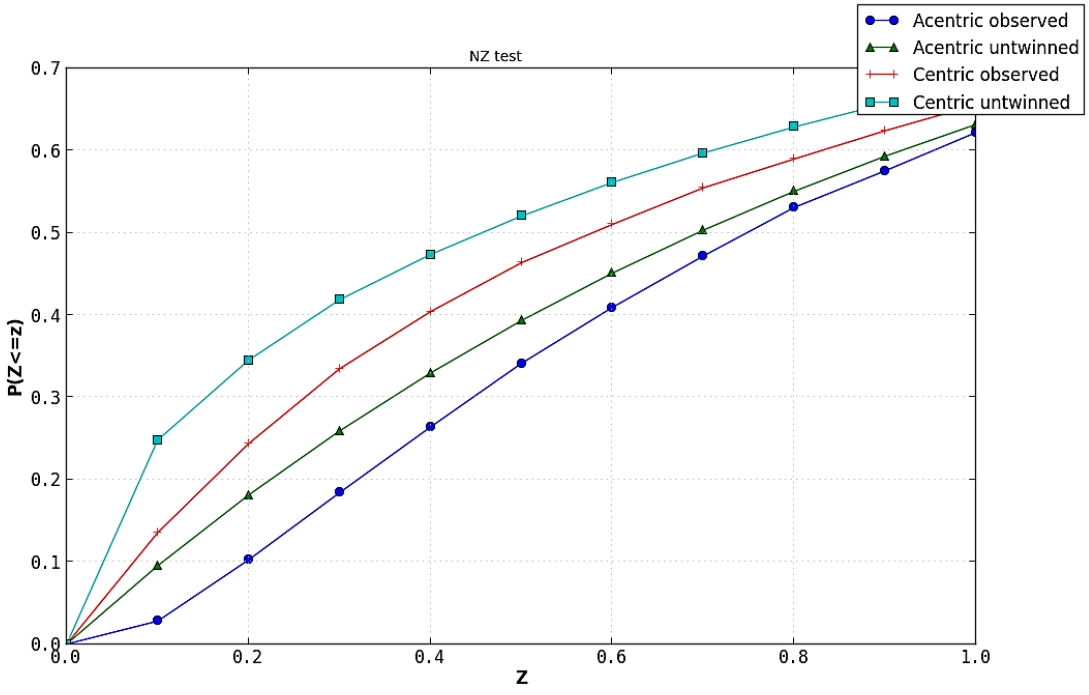


Figure 3.31. **NZ test to detect twinning or possible translational NCS.**

The plot of Z-score vs P-value (the probability equal to 9.982e-01) to detect the presence of twinning or possible translational non-crystallographic symmetry (NCS). The sigmoidal curve of the plot indicates the presence of twinning or possible NCS.

The structure of *H. pylori* FliS showed the presence of antiparallel four-helix bundle (Figure 3.32). The residues 3-20 of N-terminal region were disordered hence they were removed from the final structure. The C-terminal 5 residues were also removed. The final structure contained residues 21-121. The N-terminal conserved residues (for example Tyr 10, Tyr 28) plays a very important role in flagellin binding (Figure 3.32). It has been observed in *A. aeolicus* that two out of the four protomers from the asymmetric unit in the FliS crystal have their N-termini disordered while other two molecules showed the presence of a quasi-helical cap formed by the 15 N-terminal residues (Evdokimov *et al*., 2003).

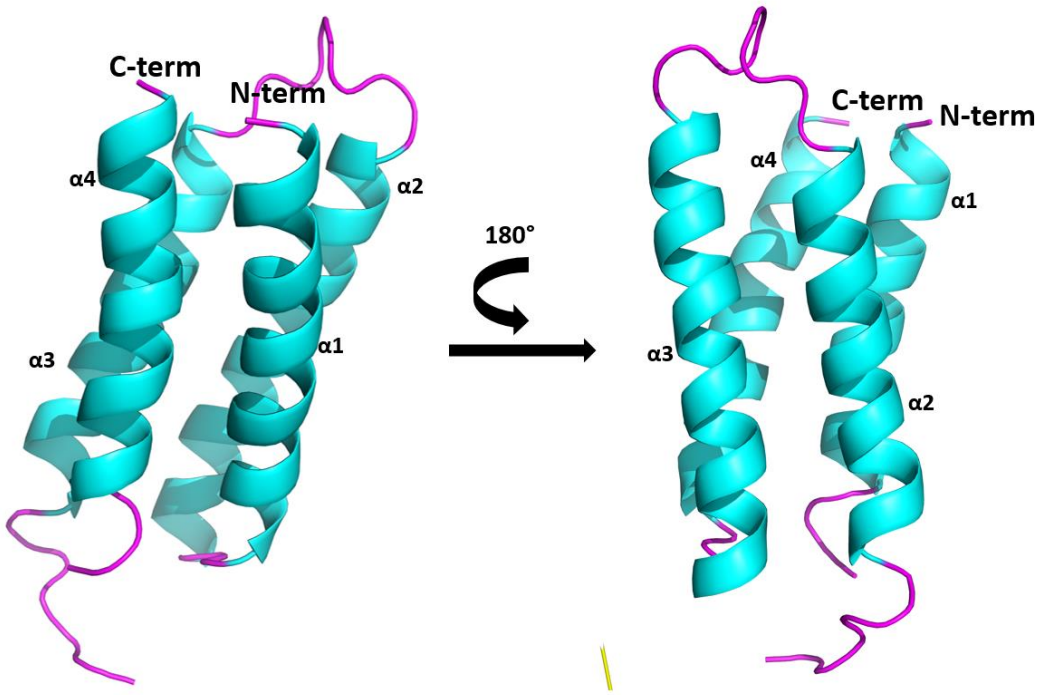


Figure 3.32. **The structure of *H. pylori* FliS (21-121).**

Two orientations of the *H. pylori* FliS structure, related by 180° rotations are drawn using PyMOL software. An antiparallel four-helix bundle is shown in cyan color and the linker segments are in magenta color.

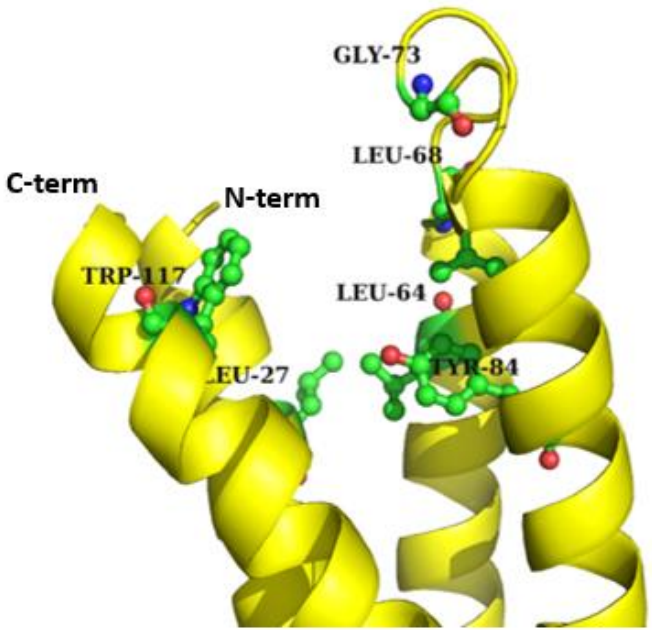


Figure 3.33. **The structure of N-terminal conserved residues in FliS.**

The core helix bundle of FliS in shown is yellow. The highly conserved residues which are important in FliS-flagellin interaction are shown in green. The moderately conserved residues are not shown in this structure. The N-terminal 19 residues were removed from the final structure. Tyr 10 which is one of the highly conserved residues, is not shown. The structure was generated using PyMOL software.

Table 3.3. **Data collection and refinement statistics**

| | |
|---|---|
| Resolution range (Å) | 46.13-2.45 (2.74-2.45) |
| Space group | P $2_1$ $2_1$ $2_1$ |
| Unit cell | a=33.86, b=82.63, c=92.26<br>α=β=γ=90 |
| Observed reflections | 70258 |
| Unique reflections | 10092 (2777) |
| Redundancy | 6.96 (7.15) |
| Completeness (%) | 99.97 (100) |
| Wilson B-factor | 28.34 |
| $R_m$ | 0.148 (0.214) |
| <I/σ(I)> | 8.2 (7.9) |
| **Details of refinement** | |
| R-work | 0.4028 (0.3728) |
| R-free | 0.4693 (0.4233) |
| Number of non-hydrogen atoms | 836 |
| water | 0 |
| RMS(bonds) | 0.005 |
| RMS(angles) | 0.835 |
| Ramachandran favored (%) | 91.3 |
| Ramachandran outliers (%) | 2.9 |
| Average B-factor | 32.30 |

(Values in parentheses are for the highest resolution shell.)

*Rmerge= ΣhklΣi|Ii(hkl) - <I(hkl)>|/ ΣhklΣi Ii(hkl)

Footnote- The data showed reflections upto 2.1 Å but for data processing we used reflections only upto 2.45 Å since beyond 2.45 Å, the data was not complete.

### 3.5.6 FliS: FlaA complex

Crystal structure of FliS in complex with flagellin from *A. aeolicus* was already determined (Evdokimov *et al*., 2003). The FliS from *A. aeolicus* shows 31% sequence identity with *H. pylori* FliS (sequence data not shown). It has been observed that *Salmonella* FliS interacts with the C-terminus of FliC (Auvray *et al*., 2001; Ozin *et al*., 2003), hence we decided to check this interaction in *H. pylori*. GST-pulldown method was used to check the interaction between FliS and N-terminus of flagellin. Both the proteins were mixed in approximately 1:2 molar ratios. 100 µL of 70% GS beads were incubated with GST-$FlaA_2$ (200 µg of GST-$FlaA_2$) at room temperature for 10 mins. The sample was then washed with 400 µL of buffer B (10 mM Tris pH 7.5, 150 mM NaCl, 2 mM DTT and 0.5 mM EDTA). The washing step was repeated 3 more times. 25 µL of the sample was mixed with 200 µg of FliS and incubated for another 20 minutes at room temperature. The incubated protein mixture was again washed 4 times with buffer B to remove unbound protein. 5 µL of the sample was mixed with 5 µL of loading buffer and run on SDS-PAGE gel. Figure 3.33 shows that GST-FlaA binds to *H. pylori* FliS.

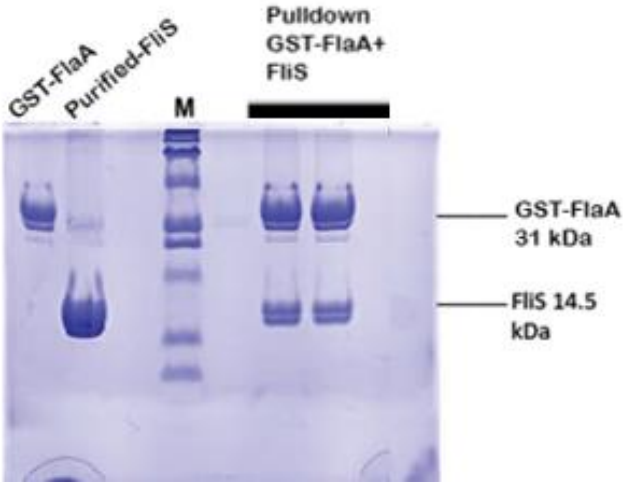


Figure 3.34. **FliS-FlaA complex pulldown.**

100 µL of 70% GS beads were incubated with GST-$FlaA_2$ (200 µg of GST-$FlaA_2$) at room temperature for 10 mins. The sample was then washed with 400 µL of buffer B (10 mM Tris pH 7.5, 150 mM NaCl, 2 mM DTT and 0.5 mM EDTA). The washing step was repeated 3 more times. 25 µL of the sample was mixed with 200 µg of FliS and incubated for another 20 minutes at room temperature. The washing step was repeated four times to remove unbound protein. 5 µL of the sample was mixed with 5 µL of loading buffer and run on 15% SDS-PAGE gel.

## 4. Discussion

### 4.1 Protein purification and crystallization of FliH, FliI and HP0256

This project started with the *in-vitro* characterization of two constructs of FliH (FliH 57-258 and FliH 73-258). When this study began, there was no structure available for *Salmonella* FliH/ *H. pylori* FliH. The genes for both the proteins were cloned successfully and the protein purification was done using GS-affinity chromatography. These proteins were subjected to calibrated SEC to determine the approximate molecular masses. The molecular weight of FliH (57-258) from its sequence is 23 kDa while the predicted molecular weight from SEC was estimated to be 60kDa. Similarly, the molecular weight of FliH (73-258) from its sequence is 21 kDa while the predicted molecular weight from SEC was estimated to be 60 kDa. The SEC results (the higher predicted molecular weight as well as broad peak size) suggests that both the FliH proteins form an elongated dimer in solution. The study in *Salmonella* has confirmed that FliH forms a stable dimer in solution (Minamino and Macnab, 2000). Further experiments are necessary to confirm the oligomerization state of this protein.

The crystallization experiments of FliH were performed successfully. For FliH (57-258), the major issue was the reproducibility of the crystals. Even after extensive optimization, it was difficult to reproduce crystals. For FliH (73-258), crystal reproducibility was not an issue but the diffraction quality of the crystals was poor. The crystals of FliH (73-258) showed weak diffraction pattern and highly anisotropic hence we were not able to determine the structure of FliH. Although low/moderate anisotropy can be tackled by some data reduction software like HKL-2000 but processing of highly anisotropic data becomes challenging (Wlodawer *et al*., 2014). Further crystallization studies are required for both these proteins. Reductive methylation of proteins has been shown to improve the quality of the crystals in some cases (Walter *et al*., 2006), therefore it can be used as one of the strategies to improve quality of these crystals in future experiments. Slightly longer constructs of FliH may be used in future experiments to check crystallization. Another approach to improve crystallization is by leaving the tag on the proteins (His-tag). It has been observed for many proteins that a small tag like His-tag helps in crystallization of protein.

Previously in Dr. Moore's lab, full length FliI as well as an N-terminal constructs of FliI have been cloned, purified and characterized. Due to poor solubility of the FliI, limited studies

were carried out on structural characterization of this protein. It is known about the role of glutamate 193 in ATP hydrolysis, hence we decided to clone the FliIE193Q mutant. The carboxyl group of glutamate 193 polarizes a water molecule for the nucleophilic attack to the γ-phosphate of ATP (a general base catalysis). Therefore, the replacement of Glutamate by Glutamine should result in the complete loss of ATPase activity (Shimabukuro *et al.,* 2003; Kazetani *et al.,* 2009). The idea behind the expression of *H. pylori* FliI E/Q mutant in the presence of AMP-PNP (a non-hydrolysable analog of ATP) was to reveal whether this substrate binds to the active site of the FliI which might create a conformational change thus regaining ATPase activity allosterically. In presence of AMP-PNP, the yield of complete hexameric FliI rings increases by 20-fold, presumably by stabilizing the assembled oligomers (Claret *et al.,* 2003).

FliIE193Q mutant was cloned and purified successfully. The yield of FliIE193Q mutant protein was low similar to wild type FliI. The FliIE193Q mutant precipitated at every stage of purification and continued till it reaches to the concentration of 2 mg/mL. From the gel filtration results for *H. pylori* FliIE193Q mutant, it was predicted to be a monomer in solution. The theoretical molecular mass of FliIE193Q mutant from its sequence was 47 kDa. The predicted molecular mass of FliI E/Q mutant from gel filtration chromatography was found to be 65 kDa. Despite conditions presumably favorable for hexamerization (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 5 mM $MgCl_2$, 100 mM Urea, 10 mM AMP-PNP and 1 mM DTT), FliI E193Q mutant did not form hexamers. This can be because of the experimental conditions. Claret and colleagues have observed FliI ring formation by electron microscopy (Claret *et al*., 2003). Furthermore, they have observed by gel filtration chromatography that the FliI forms higher order oligomers/hexamers in the presence of ATP however 60% of the total protein remains as a monomer. Hence it is possible that in this project, even if FliI is forming hexamer in very low concentration, it is impossible to detect by gel filtration. Electron microscopy can be used to observe hexamerization of FliI. Another possible reason for absence of hexamers could be that the protein is not folded properly. The partial misfolded protein may not be able to form hexamers. Another possible reason for absence of FliI hexamer could be that FliI needs another protein/chaperone to form hexamer. Recently it has been observed that FliJ promotes FliI hexamerization by binding at the center of FliI hexameric ring structure (Ibuki *et al*, 2013). Hence further experimental evidence is required to confirm the oligomerization state of FliI in *H. pylori*. The crystallization experiments of the FliIE193Q mutant protein did not produce any crystals.

This can be due to precipitation issues observed during purification or probable misfolding of FliI.

FliJ plays a very important role in the flagellar export system as lack of FliJ results in leaky motile phenotype in *Salmonella* (Minamino *et al.*, 2000). FliJ binds and promotes cycling of the export chaperones FlgN and FliT (Douillard *et al*., 2010). It interacts with FliH, FliI as well as with export gate component FlhA (Minamino & Macnab, 2000). It is a general component of flagellar export apparatus and has chaperone-like activity. Recent bioinformatics studies have identified HP0256 from *H. pylori* as a potential FliJ homologue (Douillard *et al*., 2010). FliJ promotes FliI hexamerization by binding at the center of FliI hexameric ring structure (Ibuki *et al*, 2013). Since FliJ acts as the adapter that links the substrate-chaperone complexes to export gate, the structural details of HP0256 will reveal the insights of the export process. The restoration of motility has been observed after complementing the gene in HP0256 knock out mutant lacking motility (Douillard *et al*., 2010). To prove that HP0256 is a FliJ homolog we decided to characterize HP0256. After successful cloning and protein purification of HP0256, SEC was carried out. The molecular weight of HP0256 from its sequence is 16 kDa while the estimated molecular mass from SEC was 24 kDa hence it can be presumed that FliJ remains as a monomer in solution. Unfortunately, no crystals were observed for HP0256 using standard crystallization conditions (commercially available crystallization kits - approximately over 500 conditions). Other approaches like Y2H genetic analysis, gene complementation studies cn be used to confirm that HP0256 is a FliJ homolog.

In summary, we were successful in cloning and purification of FliI, HP0256 and crystallization of FliH constructs.

**4.2 FliH:FliI E193Q binding studies**

The SEC was used to study the interaction between FliH and FliIE193Q mutant. Based on SEC results the estimated molecular mass of the FliH:FliIE193Q complex was found to be 75 kDa. Molecular weights of the individual proteins (FliH and FliIE193Q) were 21kDa and 47kDa respectively. Hence there is possibility of $FliH_2$:FliI complex but just based on SEC results it is difficult to determine the exact stoichiometry of the complex. Further experiments are required to confirm the exact molar ratio of the FliH and FliI in the complex.

### 4.3 Crystallization and the structure of FliS

Purification and crystallization of FliS was carried out successfully. The FliS crystals diffracted to 2.1 Å resolution but the data was incomplete beyond 2.45 Å resolution therefore for solving the structure of FliS we used data upto 2.45 Å resolution. The structure is not fully refined due to the data quality. The data quality also suggested that there is possibility of twinned crystal therefore to get good data, we might attempt to re-crystallize this protein in future. The structure of FliS showed an antiparallel four-helix bundle. The N-terminal residues were disordered in the structure hence were removed from the final structure. Apart from Tyr10, all other highly conserved residues were present in the final structure. The N-terminal conserved residues form a hydrophobic pocket for the binding of a flagellin molecule.

The GST-pulldown results confirm that, FliS interacts with the flagellin protein in *H. pylori*.

## 5. Future directions

Crystal screens of FliH so far have not yielded any structure. Thus for structural studies of *H. pylori* FliH, slightly longer constructs might give good diffraction data. Also, reductive methylation of lysine residues can be a good strategy to improve crystal quality. Crystallization of FliIE193Q mutant and HP0256 require further experiments. We have shown in this thesis that FliH and FliIE193Q forms a complex, sophisticated techniques like MALS will be useful to confirm the oligomerization state of the complex. The structure of FliS-flagellin complex will be interesting to reveal substrate-chaperone interaction in *H. pylori*.

## 6. Bibliography


Abby, S. and Rocha, E. (2012). The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host cell adapted systems. PLoS Genet. 8, e1002983.

Abrusci, P., McDowell, M., Lea, S., Johnson, S. (2014). Building a secreting nanomachine: a structural overview of the T3SS. Curr Opin Struct Biol. *25*,111–117

Abrusci, P., Vergara-Irigaray, M., Johnson, S., Beeby, M., Hendrixon, D., Roversi, P., Friede, M., Deane, J., Jensen, G., Tang, C., Lea, S. (2013). Architecture of the major component of the type III secretion system export apparatus. Nat Struct Mol Biol. *20,* 99–104.

Albertini, A., Caramori, T., Crabb, W., Scoffone, F., Galizzi, A. (1991). The flhA locus of *Bacillus subtilis* is part of a large operon coding for flagellar structures, motility functions, and an ATPase-like polypeptide. J Bacteriol *173,* 3573-3589.

Aldridge, P., and Hughes, K. (2002). Regulation of flagellar assembly. Curr Opin Microbiol. *5*, 160-165.

Allan, E., Dorrell, N., Foynes, S., Anyim, M., Wren, B. W. (2000). Mutational analysis of genes encoding the early flagellar components of *Helicobacter pylori*: evidence for transcriptional regulation of flagellin A biosynthesis. J. Bacteriol. *182*, 5274–5277.

Allaoui, A., Woestyn, S., Sluiters, C., Cornelis, G. R. (1994). YscU, a *Yersinia enterocolitica* inner membrane protein involved in Yop secretion. J Bacteriol. *176*, 4534–4542.

Alm, R., Ling, L., Moir, D., King, B., Brown, E., Doig, P., Smith, D., Noonan, B., Guild, B., deJonge, , Carmel, G., Tummino, P., Caruso, A., Uria-Nickelsen, M., Mills, D., Ives, C., Gibson, R., Merberg, D., Mills, S., Jiang, Q., Taylor, D., Vovis, G., Trust, T., (1999). Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature. *397,* 176–180.

Alvarez-Martinez, C. and Christie P. (2009). Biological diversity of prokaryotic type IV secretion systems. Microbiol Mol Biol Rev. *73,* 775–808.

Asherie, N. (2004). Protein crystallization and phase diagrams. Methods *34,* 266–272.

Auvray, F., Ozin, A. J., Claret, L., Hughes, C. (2002). Intrinsic membrane targeting of the flagellar export ATPase FliI: Interaction with acidic phospholipids and FliH. J Mol Biol. *318,* 941-950.

Auvray, F., Thomas, J., Fraser, G., Hughes, C., (2001). Flagellin polymerization control by a cytosolic export chaperone. J Mol Biol. *308,* 221-9.

Bai, F., Morimoto, Y., Yoshimura, S., Hara, N., Kami-Ike, N., Namba K., Minamino, T. (2014). Assembly dynamics and the roles of FliI ATPase of the bacterial flagellar export apparatus. Sci Rep *4,* 06528.

Bange, G., Kümmerer, N., Engel, C., Bozkurt, G., Wild, K., Sinning, I. (2010). FlhA provides the adaptor for coordinated delivery of late flagella building blocks to the type III secretion system. PNAS *107*, 11295–11300.

Beier, D., Spohn, G., Rappuoli, R., and Scarlato V. (1997). Identification and characterization of an operon of *Helicobacter pylori* that is involved in motility and stress adaptation. J. Bacteriol. *179*, 4676-4683.

Bennett, J. and Hughes, C. (2000). From flagellum assembly to virulence: the extended family of type III export chaperones. Trends Microbiol. *8*, 202-204.

Berg, H. and Turner L. (1979). Movement of microorganisms in viscous environments. Nature. *278*, 349–351.

Berg, J., Tymoczko, J., & Stryer, L. (2002). A Rotory Motor Drives Bacterial Motion. Biochemistry, (New York: W H Freeman), pp. 993.

Bergfors, T. (2003). Seeds to crystals. J Struct Biol. *142*, 66–76.

Bernal, R. and Stock, D. (2004). Three-dimensional structure of the intact Thermus thermophilus H+-ATPase/synthase by electron microscopy. Structure. *12*, 1789–1798.

Blaser, M. (2006). Who are we? Indigenous microbes and the ecology of human diseases. EMBO Rep. *7*, 956–960.

Blaser, M. and Atherton, J. (2004). Review *Helicobacter pylori* persistence: biology and disease. J Clin Invest. *113*, 321-333.

Blocker, A., Komoriya, K., Aizawa, S. (2003). Type III secretion systems and bacterial flagella: insights into their function from structural similarities. PNAS USA. *100*, 3027-3030.

Bonis, M., Ecobichon, C., Guadagnini, S., Prevost, M. C., Boneca, I. (2010). A M23B family metallopeptidase of *Helicobacter pylori* required for cell shape, pole formation and virulence. Mol Microbiol. *78*, 809–19.

Bourret, R. and Stock, A. (2002). Molecular Information Processing: Lessons from Bacterial Chemotaxis. J Biol Chem. *277*, 9625-9628.

Bren, A. and Eisenbach, M. (2000). How Signals Are Heard during Bacterial Chemotaxis: Protein-Protein Interactions in Sensory Signal Propagation. J Bacteriol. *182*, 6865-6873.

Brymora, A., Valova, V., Robinson, P. (2004). Protein-protein interactions identified by pull-down experiments and mass spectrometry. Curr Protoc Cell Biol. *17*, 1-51.

Buttner, D. (2012). Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. Microbiol. Mol. Biol. Rev. *76*, 262–310.

Buttner D. and He, S. (2009). Type III protein secretion in plant pathogenic bacteria. Plant Physiol. *150,* 1656–64.

Chatterjee, S., Chaudhury, S., McShan, A. C., Kaur, K., De Guzman R. (2013). Structure and biophysics of type III secretion in bacteria. Biochemistry. *52,* 2508–2517.

Chayen, N. (1998). Comparative studies of protein crystallization by vapour diffusion and microbatch techniques. Acta Crystallogr D Biol Crystallogr. *54,* 8–15.

Chayen, N. (2004).Turning protein crystallization from an art into a science. Curr Opin Struct Biol. *14,* 577–583.

Chen, L., and Helmann, J. (1994) The Bacillus subtilis sigma D-dependent operon encoding the flagellar proteins FliD, FliS, and FliT. J. Bacteriol. *176,* 3093–3101.

Chen, S., Beeby, M., Murphy, E., Leadbetter, J., Hendrixson, D., Briegel, A., Li, Z., Shi, J., Tocheva, E., Mu¨ller, A., Dobro, M., Jensen, G. (2011). Structural diversity of bacterial flagellar motors. The EMBO Journal, 1–10.

Chevance, F. and Hughes, K. (2008). Coordinating assembly of a bacterial macromolecular machine. Nat. Rev. Microbiol. *6,* 455–465.

Chilcott, G. and Hughes, K. (2000). Coupling of flagellar gene expression to flagellar assembly in *Salmonella enterica serovar typhimurium* and *Escherichia coli*. Microbiol Mol Biol Rev. *64,* 694-708.

Claret, L., Calder, S., Higgins, M., and Huges, C. (2003). Oligomerization and activation of the FliI ATPase central to the bacterial flagellum assembly. Mol. Microbiol. *48,* 1349–1355.

Colland, F., Rain, J., Gounon, P., Labigne, A., Legrain, P. De Reuse, H. (2001). Identification of the *Helicobacter pylori* anti-sigma 28 factor. Mol Microbiol. 41, 477-487.

Collazo, C. and Galán, J. (1996). Requirement for exported proteins in secretion through the invasion-associated type III system of *Salmonella typhimurium*. Infect Immun. *64,* 3524–3531.

Collazo, C., Galán, J. (1997). The invasion-associated type III system of *Salmonella typhimurium* directs the translocation of Sip proteins into the host cell. Mol Microbiol. *24,* 747–756.

Collazo, C., Zierler, M., Galán, J. (1995). Functional analysis of the *Salmonella typhimurium* invasion genes invI and invJ and identification of a target of the protein secretion apparatus encoded in the inv locus. Mol Microbiol. *15,* 25–38.

Cordes, F., Komoriya, K., Larquet, E., Yang, S., Egelman, E., Blocker, A., Lea, S. (2003). Helical structure of the needle of the type III secretion system of *Shigella flexneri*. J Biol Chem. *278,* 17103–17107.

Cordingley, M., Callahan, P., Sardana, V., Garsky, V., Colonno, R. (1990). Substrate requirements of human rhinovirus 3C protease for peptide cleavage *in vitro*. J Biol Chem. *265,* 9062-9065.

Cornelis, G. (2006). The type III secretion injectisome. Nat Rev Microbiol. *4*, 811–825.

Correa, V., Majerczak, D., Ammar, el-D., Merighi, M., Pratt, R., Hogenhout, S., Coplin, D., Redinbaugh, M. (2012). The bacterium *Pantoea stewartii* uses two different type III secretion systems to colonize its plant host and insect vector. Appl Environ Microbiol. *78*, 6327–6336.

Cross, R. and Muller, V. (2004). The evolution of A-, F-, and V-type ATP synthases and ATPases: reversals in function and changes in the H+/ATP coupling ratio. FEBS Lett. *576*, 1–4.

Dickson, V., Silvester, J., Fearnley, I., Leslie, A., Walker, J. (2006). On the structure of the stator of the mitochondrial ATP synthase. EMBO J. *25*, 2911–2918.

Diepold, A. & Armitrage, J. (2015). Type III secretion systems: the bacterial flagellum and the injectisome. Phil. Trans. R.Soc. B. *370*. 1-19.

Douillard, F., Ryan, K., Caly, D., Hinds, J., Witney, A., Husain, S., and O'Toole, P. (2008). Posttranscriptional Regulation of Flagellin Synthesis in *Helicobacter pylori* by the RpoN Chaperone HP0958. J. Bacteriol. *190*, 7975–84.

Douillard, F., Ryan, K., Lane, M., Caly, D., Moore, S., Penn, C., Hinds, J., O'Toole, P. (2010). The HP0256 gene product is involved in motility and cell envelope architecture of *Helicobacter pylori*. BMC Micro. *10*, 106-113.

Dunn, S.D., McLachlin, D.T., Revington, M. (2000). The second stalk of *Escherichia coli* ATP synthase. Biochim Biophys Acta. *1458*, 356–363.

Eaton, K., Morgan, D., Krakowka, S. (1989). *Campylobacter pylori* virulence factors in gnotobiotic piglets. Infect Immun. *57*, 1119-1125.

Eaton, K., Suerbaum, S., Josenhans, C., Krakowka, S. (1996). Colonization of gnotobiotic piglets by *Helicobacter pylori* deficient in two flagellin genes. Infect Immun. *64*, 2445–8.

Eichelberg, K., Ginocchio, C., Galan, J. (1994). Molecular and functional characterization of the *Salmonella typhimurium* invasion genes invB and invC: Homology of InvC to the $F_OF_1$ ATPase family of proteins. J. Bacteriol. *176*, 4501–10.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486-501.

Epier C, Dickenson N, Bullitt E, Picking W. (2012). Ultrastructural analysis of IpaD at the tip of the nascent MxiH type III secretion apparatus of *Shigella flexneri*. J. Mol. Biol. *420*, 29–39.

Erhardt, M., Namba, K., Hughes, K. (2010). Bacterial nanomachines: The flagellum and Type III injectisome. Cold Spring Harb Perspect Biol. *2*, a000299.

Evans, L., Stafford, G., Ahmed, S., Fraser, G., Hughes, C. (2006). An escort mechanism for cycling of export chaperones during flagellum assembly. PNAS USA *103*, 17474–17479.

Evdokimov, A., Phan, J., Tropea, J., Routzahn, K., Peters, H., Pokross, M. Waugh, D. (2003). Similar modes of polypeptide recognition by export chaperones in flagellar biosynthesis and type III secretion. Nature Struct. Biol. *10,* 789–793.

Fan, F. and Macnab, R. (1996). Enzymatic characterization of FliI: an ATPase involved in flagellar assembly in *Salmonella typhimurium*. J. Biol. Chem. *271,* 31981-31988.

Fan, F., Ohnishi, K., Francis, N., Macnab, R. (1997). The FliP and FliR proteins of *Salmonella typhimurium*, putative components of the type III flagellar export apparatus, are located in the flagellar basal body. Mol Microbiol. *26,* 1035–1046.

Feniouk, B., Suzuki, T., Yoshida, M. (2007). Regulatory interplay between proton motive force, ADP, phosphate, and subunit epsilon in bacterial ATP synthase. J. Biol. Chem. *282,* 764– 772.

Ferlay, J., Shin, H., Bray, F., Forman, D., Mathers, C., Parkin, D. (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer. *127,* 2893–2917.

Ferrero, R. and Lee, A. (1988). Motility of *Campylobacter jejuni* in a viscous environment: comparison with conventional rod-shaped bacteria. J Gen Microbiol. *134,* 53–59.

Ferris, H. and Minamino, T. (2006). Flipping the switch: bringing order to flagellar assembly. Trends Microbiol. *14,* 519-526.

Fields, K., Plano, G. V., Straley, S. (1994). A low-Ca2+ response (LCR) secretion (ysc) locus lies within the lcrB region of the LCR plasmid in *Yersinia pestis*. J Bacteriol. *176,* 569–579.

Forgac, M. (2007). Vacuolar ATPases: rotary proton pumps in physiology and pathophysiology. Nat Rev Mol Cell Biol. *8,* 917–929.

Fraser, G., Bennett, J., Hughes, C. (1999). Substrate-specific binding of hook-associated proteins by FlgN and FliT, putative chaperones for flagellum assembly. Mol Microbiol. *32,* 569-580.

Fraser, G., Gonzalez-Pedrajo, B., Tame, J., Macnab, R. (2003). Interactions of FliJ with the *Salmonella* type III flagellar export apparatus. J. Bacteriol. *185,* 5546–5554.

Galàn, J. (2009). Common themes in the design and function of bacterial effectors. Cell Host Microbe. *5,* 571–579.

Galàn, J. and Wolf-Watz H. (2006). Protein delivery into eukaryotic cells by type III secretion machines. Nature. *444,* 567–573.

Galán, J., Ginocchio, C., Costeas, P. (1992). Molecular and functional characterization of the *Salmonella* invasion gene invA: homology of InvA to members of a new protein family. J Bacteriol. *174,* 4338–4349.

Galàn, J., Lara-Tejero, M., Marlovits, T., Wagner, S. (2014). Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells. Annu Rev Microbiol. *68,* 415–438.

Galkin, V., Schmied, W., Schriadt, O., Marlovits, T., Egelman, E. (2010). The structure of the *Salmonella typhimurium* type III secretion system needle shows divergence from the flagellar system. J Mol Biol. *396,* 1392–1397.

Gao, B., Lara-Tejero, M., Lafebre, M., Goodman, A., Galàn, J. (2014). Novel components of flagellar system in epsilonproteobacteria. MBio. *5,* e01349-14.

Ginocchio, C. and Galán, J. (1995). Functional conservation among members of the *Salmonella typhimurium* InvA family of proteins. Infect Immun. *63,* 729–732.

Gonzàlez-Pedrajo, B., Fraser, G., Minamino, T. and Macnab, R. (2002). Molecular dissection of *Salmonella* FliH, a regulator of the ATPase FliI and the type III flagellar protein export pathway. Mol Microbiol. *45*, 967-982.

Gonzàlez-Pedrajo, B., Minamino, T., Kihara, M., Namba, K. (2006). Interactions between C ring proteins and export apparatus components: a possible mechanism for facilitating type III protein export. Mol Microbiol. *60*, 984-998.

Groisman, E. and Ochman, H. (1993). Cognate gene clusters govern invasion of host epithelial cells by *Salmonella typhimurium* and *Shigella flexneri*. EMBO J. *12,* 3779–3787.

Harper, S. and Speicher, D. (2011). Purification of proteins fused to glutathione S-transferase. Methods Mol Biol. *681,* 259-280.

Hu, B., Morado, D., Margolin, W., Rohede, J., Arizmendi, O., Picking, W., Picking, W., Liu, J. (2015). Visualization of the type III secretion sorting platform of *Shigella flexneri*. PNAS. 112 1047-1052

Hueck, C. (1998). Type III protein secretion systems in bacterial pathogens of animals and plants. Mol. Biol. Rev. *62*, 379–433.

Ibuki, T., Uchida, Y., Hironaka, Y., Namba, K., Imada, K., Minamino, T. (2013). Interaction between FliJ and FlhA, III Export Apparatus Components of the Bacterial Flagellar Type. J Bacteriol. *195*, 466-73.

Ibuki. T., Imada. K., Minamino. T., Kato. T., Miyata. T., Namba. K., (2011). Common architecture of the flagellar type III protein export apparatus and F- and V-type ATPases. Nat Struct Mol Biol *18,* 277–282.

Ibuki. T., Shimada, M., Minamino, T., Namba, K., Imada, K. (2008). Crystallization and preliminary X-ray analysis of FliJ, a cytoplasmic component of the flagellar type III protein-export apparatus from *Salmonella* sp. Acta Cryst. *F65*, 47-50.

Iino, R., Noji, H. (2013). Operation mechanism of $F_O$ $F_1$-adenosine triphosphate synthase revealed by its structure and dynamics. IUBMB Life. *65*, 238-246.

Imada, K., Minamino. T., Uchida, Y., Kinoshita, M., Namba. K. (2016). Insight into the flagella type III export revealed by the complex structure of the type III ATPase and its regulator. PNAS. *113*, 3633-3638.

Imada. K., Minamino. T., Tahara. A., Namba. K. (2007). Structural similarity between the flagellar type III ATPase FliI and F1-ATPase subunits. Proc Natl Acad Sci USA *104*, 485–490.

Imae, Y. and Atsumi, T. (1989). Na+-driven bacterial flagellar motors. J Bioenerg Biomembr. *21*, 705-716.

Jefferies, K., Cipriano, D., Forgac, M. (2008) Function, structure and regulation of the vacuolar (H+)-ATPases. Arch. Biochem. Biophys. *476*, 33–42.

Jenks, P., Foynes, S., Ward, S., Constantinidou, C., Penn, C., Wren, B. (1997). A flagellar-specific ATPase (FliI) is necessary for flagellar export in *Helicobacter pylori*. FEMS Microbiol Lett. *152*, 205 -211.

Josenhans, C., and Suerbaum, S. (2002). The role of motility as a virulence factor in bacteria. Int J Med Microbiol. *291*, 605–614.

Kabsch, W. (2010). *XDS.* Acta Cryst. D66, 125-132.

Kawagishi, I., Müller, V., Williams, A., Irikura, V., and Macnab, R. (1992). Subdivision of flagellar region III of the *Escherichia coli* and *Salmonella typhimurium* chromosomes and identification of two additional flagellar genes. J. Gen. Microbiol. *138,* 1051–1065.

Kawamoto, A., Morimoto, Y., Miyata, T., Minamino, T., Hughes, K., Kato, T., Namba, K. (2013). Common and distinct structural features of *Salmonella injectisome* and flagellar basal body. Sci Rep 3, 3369.

Kazetani, K., Minamino, T., Miyata, T., Kato, T. & Namba, K. (2009).ATP-induced FliI hexamerization facilitates bacterial flagellar protein export. Biochem. Biophys. Res. Commun. *388*, 323–327.

Kim, J., Chang, J., Chung, S. Yum, J. (1999). Molecular cloning and characterization of the *Helicobacter pylori fliD* gene, an essential factor in flagellar structure and motility. J. Bacteriol. *181,* 6969-6976.

Kishikawa, J., Ibuki, T., Nakamura, S., Nakanishi, A., Minamino, T., Miyata, T., Namba, K., Konno, H., Ueno, H., Imada, K., Yokoyama, K. (2013). Common evolutionary origin for the rotor domain of rotary ATPases and flagellar protein export apparatus. PLoS One. *8*, e64695.

Kitagawa, N., Mazon, H., Heck, A., Wilkens, S. (2008). Stoichiometry of the peripheral stalk subunits E and G of yeast V1-ATPase determined by mass spectrometry. J Biol Chem. *283,* 3329–3337.

Kosarewicz, A., Konigsmaier, L., Marlovits, T. (2012). The blueprint of the type-3 injectisome. Philos Trans R Soc Lond B Biol Sci. *367,* 1140–1154.

Kubori, T., Matsushima, Y., Nakamura, D., Uralil, J., Lara-Tejero, M., Sukhan, A., Galan, J. & Aizawa, S.-I. (1998). Supramolecular structure of the *Salmonella typhimurium* type III protein secretion system. Science. *280*, 602–605.

Kubori, T., Sukhan, A., Aizawa, S., Galàn, J. (2000). Molecular characterization and assembly of the needle complex of the *Salmonella typhimurium* type III protein secretion system. Proc Natl Acad Sci USA. *97,* 10225–10230.

Kudryashev M, Stenta M, Schmelz S, Amstutz M, Wiesand U, et al. (2013). In situ structural analysis of the *Yersinia enterocolitica* injectisome. eLife *2*, e00792.

Kutsukake, K. (1994). Excretion of the anti-sigma factor through a flagellar substructure couples flagellar gene expression with flagellar assembly in *Salmonella typhimurium*. Mol. Gen. Genet. 243**,** 605-612.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680-685.

Lai, Y., Yang, J., Lin, T., Lin, J., Wang, J. (2006). *Helicobacter pylori* infection and CagA protein translocation in human primary gastric epithelial cell culture. Helicobacter. *11*, 451–459.

Lam, W., Woo, E., Kotaka M., Tam, W., Leung, Y., Ling, T., Au, S. (2010). Molecular interaction of flagellar export chaperone FliS and cochaperone HP1076 in *Helicobacter pylori.* FASEB journal, *24*, 4020-4032.

Lane, M., O'Toole, P. Moore, S. (2006). Molecular Basis of the Interaction between the Flagellar Export Proteins FliI and FliH from *Helicobacter pylori*. J Biol Chem. *281*, 508–17.

Lee, L., Stewart, A., Donohoe, M., Bernal, R., Stock, D. (2010). The structure of the peripheral stalk of *Thermus thermophilus* H+-ATPase/ synthase. Nat Struct Mol Biol. *17,* 373–8.

Lertsethtakarn, P., Ottemann, K., Hendrixson, D. (2011). Motility and Chemotaxis in *Campylobacter* and *Helicobacter*. Annu Rev Microbiol. *65,* 389–410.

Leying, H., Suerbaum, S., Geis, G., Haas, R. (1992). Cloning and genetic characterization of a *Helicobacter pylori* flagellin gene. Mol. Microbiol. *6*, 2863-2874.

Li, S. and Kane, P. (2009). The yeast lysosome-like vacuole: endpoint and crossroads. Biochim Biophys Acta. *1793,* 650–663.

Loquet, A., Sgourakis, N., Gupta, R., Giller, K., Riedel, D., Goosmann, C., Kolbe, M., Baker, D., Becker, D., Becker, S., Lange, A. (2012). Atomic model of the type III secretion system needle. Nature *486,* 276–279.

Lorenzini. E., Singer. A., Singh. B., Lam. R., Skarina. T., Chirgadze. N., Savchenko. A., Gupta. R. (2010). Structure and protein-protein interaction studies on Chlamydia trachomatis protein CT670 (YscO Homolog). J Bacteriol *192,* 2746–2756.

Macnab, R. (1992). Genetics and biogenesis of bacterial flagella. Ann. Rev. Genet. *26,* 131-158.

Macnab, R. (1996). In *Escherichia coli* and *Salmonella typhimurium*: Cell Mol Biol. 123-145.

Macnab, R. (2003). How bacteria assemble flagella. Annu. Rev. Microbiol. *57,* 77–100.

Manson, M., Tedesco, P., Berg, H., Harold, F., Van der Drift, C. (1977). A protonmotive force drives bacterial flagella. Proc. Natl Acad. Sci USA. *74,* 3060–3064.

Marlovits, T., Kubori, T., Lara-Tejero, M., Thomas, D., Unger, V. M., Galán, J. (2006). Assembly of the inner rod determines needle length in the type III secretion injectisome. Nature. *441,* 637–40.

Marlovits, T., Kubori, T., Sukhan, A., Thomas, D., Galán, J., Unger, V. (2004). Structural insights into the assembly of the type III secretion needle complex. Science. *306,* 1040–1042.

McCoy, A., Grosse-Kunstleve, R., Adams, P., Winn, M., Storoni, L. , Read, R. (2007). J Appl Crystallogr. *40,* 658-674.

Meister, M., Lowe, G., Berg, H. (1987). The proton flux through the bacterial flagellar motor. Cell. *49,* 643-650.

Merz, A., So, M., Sheetz, M. (2000). Pilus retraction powers bacterial twitching motility. Nature *407,* 98–102.

Minamino, T. (2014). Protein export through the bacterial flagellar type III export pathway. Biochim Biophys Acta. *1843,* 1642–1648.

Minamino, T. and Imada, K. (2015). The bacterial flagellar motor and its structural diversity. Trends Microbiol. *23,* 267-274.

Minamino, T. and Macnab, R. M. (1999). Components of the *Salmonella* flagellar export apparatus and classification of export substrates. J Bacteriol. *181,* 1388–1394.

Minamino, T., and MacNab, R. (2000b). FliH, a soluble component of the type III flagellar export apparatus of *Salmonella*, forms a complex with FliI and inhibits its ATPase activity. Mol. Microbiol. *37,* 1494–1503.

Minamino, T., and Namba, K. (2008) Distinct roles of the FliI ATPase and proton motive force in bacterial flagellar protein export. Nature *451,* 485–488.

Minamino, T., Chu, R., Yamaguchi, S., Macnab, R. (2000). Role of FliJ inflagellar protein export in *Salmonella*. J. Bacteriol. *182*, 4207–4215.

Minamino, T., Gonzalez-Pedrajo, B., Kihara, M., Namba, K., Macnab, RM. (2003) The ATPase FliI can interact with the type III flagellar protein export apparatus in the absence of its regulator, FliH. J. Bacteriol. *185*, 3983–3988.

Minamino, T., Gonzalez-Pedrajo, B., Oosawa, K., Namba, K., Macnab, R. (2002). Structural properties of FliH, an ATPase regulatory component of the *Salmonella* type III flagellar export apparatus. J. Mol. Biol. *322*, 281-290.

Minamino, T., Imada, K., Namba, K. (2008b). Molecular motors of the bacterial flagella. Curr Opin Struct Biol. *18*, 693–701.

Minamino, T., Kazetani, K., Tahara, A., Suzuki, H., Furukawa, Y., Kihara, M. and Namba, K. (2006). Oligomerization of the bacterial flagellar ATPase FliI is controlled by its extreme N-terminal region. J Mol Biol. *360*, 510–519.

Minamino, T., Kinoshita, M., Imada, K., and Namba, K. (2012). Interaction between FliI ATPase and a flagellar chaperone FliT during bacterial flagellar protein export. Mol. Microbiol. *83*, 168–178.

Minamino, T., Kinoshita, M., Inoue, Y., Morimoto, Y., Ihara, K., Koya, S., Hara, N., Nishioka, N., Kojima, S., Homma, M., Namba, K. (2016). FliH and FliI ensure efficient energy coupling of flagellar type III protein export in *Salmonella*. Microbiology Open. *5*, 424–435.

Minamino, T., Tame, J., Namba, K., Macnab, R. (2001). Proteolytic analysis of the FliH/FliI complex, the ATPase component of the Type III flagellar export apparatus of *Salmonella*. J Mol. Biol. *312*, 1027-1036.

Minamino, T., Yoshimura, S., Morimoto, Y., González-Pedrajo, B., Kami-Ike, N., Namba, K. (2009). Roles of the extreme N-terminal region of FliH for efficient localization of the FliH-FliI complex to the bacterial flagellar type III export apparatus. Mol. Microbiol. *74*, 1471–1483.

Minamino, T., Morimoto, Y., Kinoshita, M., Aldridge, P., Namba, K. (2014). The bacterial flagellar protein export apparatus processively transports flagellar proteins even with extremely infrequent ATP hydrolysis. Sci Rep *4*, 7579.

Minamino,T., Imada, K., Namba, K. (2008a). Mechanisms of type III protein export for bacterial flagellar assembly. Mol Biosyst. *4*, 1105–1115.

Mobley, H., Mendz, G., Hazell, S. (2001). Basic Bacteriology and Culture. *Helicobacter pylori*: Physiology and Genetics. (Washington (DC): ASM Press), pp 2444.

Moore, S., and Jia, Y. (2010) Structure of the Cytoplasmic domain of the Flagellar secretion apparatus component FlhA from *helicobacter pylori*. J Biol Chem. *285*, 21060 –21069.

Morita-Ishihara, T., Ogawa, M., Sagara, H., Yoshida, M., Katayama, E., Sasakawa, C. (2006). *Shigella* Spa33 is an essential C-ring component of type III secretion machinery. J Biol Chem. *281,* 599–607.

Mueller, C., Broz, P., Müller, S., Ringler, P., Erne-Brand, F., Sorg, I., Kuhn, M., Engel, A., Cornelis, G. (2005). The V-antigen of Yersinia forms a distinct structure at the tip of injectisome needles. Science. *310,* 674–6.

Mulkidjanian, A., Makarova, K., Galperin, M., Koonin, E. (2007). Inventing the dynamo machine: the evolution of the F-type and V-type ATPases. Nat Rev Microbiol. *5,* 892–899.

Muskotal, A., Kiraly, R., Sebestyen, A., Gugolya, Z., Vegh, B., Vonderviszt, F. (2006). Interaction of FliS flagellar chaperone with flagellin. FEBS Lett. *580,* 3916-3920.

Nakano, M., Imamura, H., Toei, M., Tamakoshi, M., Yoshida, M., Yokoyama, K. (2008). ATP hydrolysis and synthesis of a rotary motor V-ATPase from *Thermus thermophilus*. J Biol Chem. *283,* 20789–96.

Niehus, E., Gressmann, H., Ye, F., Schlapbach, R., Dehio, M., Dehio, C., Stack, A., Meyer, T.F., Suerbaum, S., Josenhans, C., (2004). Genome-wide analysis of transcriptional hierarchy and feedback regulation in the flagellar system of *Helicobacter pylori*. Mol. Microbiol. *52,* 947–961.

Noritaka, H., Morimoto, Y., Kawamoto, A., Namba, K., Minamino, T. (2012). Interaction of the extreme N-terminal region of FliH with FlhA is required for efficient bacterial flagellar protein export. J. Bactreriol, *194*, 5353-5360.

Oot, R., Huang, L., Berry, E., Wilkens, S. (2012). Crystal Structure of the yeast vacuolar ATPase Heterotrimeric EGC (head) Peripheral Stalk Complex. Structure. *20,* 1881–1892.

Ottemann, K. and Lowenthal, A. (2002). *Helicobacter pylori* uses motility for initial colonization & to attain robust infection. Infect Immun. *70*, 1984-1990.

Ozin, A., Claret, L., Auvray, F., Hughes, C. (2003). The FliS chaperone selectively binds the disordered flagellin C-terminal D0 domain central to polymerisation. FEMS Microbiol Lett. *219,* 219-224.

Palframan, S., Kwok, T., Gabriel, K. (2012). Vacuolating cytotoxin A (VacA), a key toxin for *Helicobacter pylori* pathogenesis. Front Cell Infect Microbiol. *2,* 1-9.

Pallen, M. and Matzke, N. J. (2006). From the origin of species to the origin of bacterial flagella. Nat Rev Microbiol. *4,* 784-790.

Pallen, M., Bailey C., Beatson S. (2006). Evolutionary links between FliH/YscL-like proteins from bacterial type III secretion systems and second-stalk components of the $F_OF_1$ and vacuolar ATPases. Protein Sci. *15*, 935-941.

Paul, K., Erhardt, M., Hirano, T., Blair, D., Hughes, K. (2008). Energy source of flagellar type III secretion. Nature, *451,* 489-492.

Pellicano, R., Franceschi, F., Saracco, G., Fagoonee, S., Roccarina, D., Gasbarrini, A. (2009). *Helicobacters* and extragastric diseases. *Helicobacter*. *14,* 58-68.

Plano, G., Barve, S., Straley, S. (1991). LcrD, a membrane-bound regulator of the *Yersinia pestis* low calcium response. J. Bacteriol. *173,* 7293–7303.

Porwollik, S., Noonan, B., O'Toole, P. (1999). Molecular characterization of a flagellar export locus of *Helicobacter pylori*. Infect. Immun. *67,* 2060-2070.

Radics, J., Königsmaier, L., Marlovits, T. (2014). Structure of a pathogenic type 3 secretion system in action. Nat Struct Mol Biol. *21,* 82–87.

Rain, J. C., Selig, L., De Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schachter, V., Chemama, Y., Labigne, A., Legrain, P. (2001). The protein-protein interaction map of *Helicobacter pylori*. Nature 409**,** 211-215.

Rhee, K., Park, J., Cho, M. (2014). Helicobacter pylori: Bacterial Strategy for Incipient Stage and Persistent Colonization in Human Gastric Niches. Yonsei Med J. *55,* 1453-1466.

Roblin, P., Dewitte, F., Villeret, V., Biondi, E., Bompard, C. (2014). A *Salmonella* Type Three Secretion Effector/Chaperone Complex adopts Hexameric Ring-Like Structure. .J Bacteriol. *197,* 688-698.

Rosqvist, R., Persson, C., Hakansson, S., Nordfeldt, R., Wolf-Watz, H. (1995). Translocation of the *Yersinia* YopE and YopH virulence proteins into target cells is mediated by YopB and YopD. Contrib Microbiol Immunol. *13,* 230–34.

Rothebacher, D., Brenner, H. (2003). Burden of *Helicobacter pylori* and *H. pylori* related diseases in developed countries: recent developments and future implications. Microbes Infect *5,* 693-703.

Ryan, K., Karim, N., Worku, M., Moore, S., Penn, C., O'Toole, P. (2005). HP0958 is an essential motility gene in *Helicobacter pylori*. FEMS Microbiol Lett. *248,* 47-55.

Schoenhofen, I., Lunin, V., Julien, J., Li, Y., Ajamian, E., Matte, A., Cygler, M., Brisson, J., Aubry, A., Logan, S., Bhatia, S., Wakarchuk, W., Young, N. (2006). Structural and functional characterization of PseC, an aminotransferase involved in the biosynthesis of pseudaminic acid, an essential flagellar modification in *Helicobacter pylori*. J Biol Chem. *281,* 8907-16.

Schraidt, O. & Marlovits, T. (2011). Three-dimensional model of *Salmonella*'s needle complex at subnanometer resolution. Science. *331,* 1192–1195.

Silver, A., Kikuchi, Y., Fadl, A., Sha, J., Chopra, A., Graf, J. (2007). Interaction between innate immune cells and a bacterial type III secretion system in mutualistic and pathogenic associations. Proc Natl Acad Sci USA. *104,* 9481–9486.

Soutourina, O. and Bertin, P. (2003). Regulation cascade of flagellar expression in Gram-negative bacteria. FEMS Microbiol Rev. *27*, 505-523.

Spohn, G., and Scarlato, V. (1999). Motility of *Helicobacter pylori* is coordinately regulated by the transcriptional activator FlgR, an NtrC homolog. J Bacteriol. *181*, 593-599.

Stewart, A., Laming, E., Sobti, M., Stock, D. (2014). Rotary ATPases — dynamic molecular machines. Curr Opin Struct Biol. *25*, 40–48.

Sun, L., Dong, Y., Shi, M., Zhou, Q., Luo Z. Q., Gao, H. (2014). Two residues predominantly dictate functional difference in motility between *Shewanella oneidensis* flagellins FlaA and FlaB. J. Biol. Chem. *289,* 14547-14559.

Sycuro, L., Pincus, Z., Gutierrez, K., Biboy, J., Stern, C., Collmer, W., Salama, N. (2010). Peptidoglycan crosslinking relaxation promotes *Helicobacter pylori*'s helical shape and stomach colonization. Cell. *141*, 822–833.

The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

Thomas, D., Francis, N., Xu, C. & DeRosier, D. (2006). The three-dimensional structure of the flagellar rotor from a clockwise-locked mutant of *Salmonella enterica serovar Typhimurium*. J. Bacteriol. 188, 7039–7048.

Thomas, D., Morgan, D. & DeRosier, D. (2001). Structures of Bacterial Flagellar Motors from Two FliF-FliG Gene Fusion Mutants. J. Bacteriol. *183*, 6404–6412.

Thomas, J., Stafford, G. & Hughes, C. (2004). Docking of cytosolic chaperone-substrate complexes at the membrane ATPase during flagellar type III protein export. PNAS USA *101*, 3945–3950.

Tomb, J., White, O., Keflavage, A. *et al.*, (1997). The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature *388,* 539-47.

Trost, B. and Moore, S. (2009). Statistical characterization of the GxxxG glycine repeats in the flagellar biosynthesis protein FliH and its Type III secretion homologue YscL. BMC Micro. *9*, 72-92.

Tsang, J., and Hoover, T. (2014). Requirement of the Flagellar Protein Export Apparatus Component FliO for Optimal Expression of Flagellar Genes in *Helicobacter pylori*. J Bacteriol. *196,* 2709 –2717.

Vogler, A., Homma, M., Irikura, V., Macnab, R. (1991). *Salmonella typhimurium* mutants defective in flagellar filament regrowth and sequence similarity of FliI to, and archaebacterial ATPase subunits. J Bacteriol, *173*, 3564-72.

Wagner, S., Königsmaier, L., Lara-Tejero, M., Lefebre, M., Marlovits, T., Galàn, J. (2010). Organization and coordinated assembly of the type III secretion export apparatus. Proc Natl Acad Sci USA. *107*, 17745-50.

Walker, P., Leong, L., Ng, P., Tan, S., Waller, S., Murphy, D., Porter, A. (1994). Efficient and Rapid Affinity Purification of Proteins Using Recombinant Fusion Proteases. Nature Biotechnology *12*, 601 – 605.

Walter, T., Meier, C., Assenberg, R., Au, K., Ren, J., Verma, A., Nettleship, J., Owens, R., Stuart, D., Grimes, J. (2006). Lysine Methylation as a Routine Rescue Strategy for Protein Crystallization. Structure *14*, 1617-1622.

Warren J. and Marshall B. (1983). Unidentified curved bacilli on gastric epithelium in active chronic gastritis. Lancet. *321*, 1273–1275.

Wlodawer, A., Minor, W., Dauter, Z., Jaskolski, M. (2014). Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. FEBS J. *280*, 5705-5736.

Woestyn, S., Allaoui, A., Wattiau, P., Cornelis, G. (1994). YscN, the putative energizer of the *Yersinia* Yop secretion machinery. J Bacteriol. *176*, 1561–1569.

Wolf, E., Kim, P., and Berger, B. (1997). MultiCoil: a program for predicting two- and three-stranded coiled coils. Protein Sci. *6*, 1179–1189.

Worrall L., Lameignere, E., Strynadka, N. (2011). Structural overview of the bacterial injectisome. Curr Opin Microbiol. *14*, 3–8.

Yamaoka, Y. (2010). Mechanisms of disease: *Helicobacter pylori* virulence factors. Nat Rev Gastroenterol Hepatol. *7*, 629–641.

Yokoseki, T., Iino, T., and Kutsukake, K. (1996) Negative regulation by fliD, fliS, and fliT of the export of the flagellum-specific anti-sigma factor, FlgM, in *Salmonella typhimurium*. J. Bacteriol. *178*, 899–901.

Yokoseki, T., Kutsukake, K., Ohnishi, K., and Iino, T. (1995). Functional analysis of the flagellar genes in the fliD operon of *Salmonella typhimurium*. Microbiology *141*, 1715–1722.

Z. Otwinowski and W. Minor, " Processing of X-ray Diffraction Data Collected in Oscillation Mode ", Methods in Enzymology, Volume 276: Macromolecular Crystallography, part A, p.307-326, 1997,C.W. Carter, Jr. & R. M. Sweet, Eds., Academic Press (New York).